

GenCore version 5.1.6
Copyright (c) 1993 - 2006 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: January 25, 2006, 18:48:24 ; Search time 68.2 Seconds
(without alignments)
128.850 Million cell updates/sec

Title: US-10-797-821-1

Perfect score: 94

Sequence: 1 KSNATSYNAINIINSKVSVD 20

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 2443163 seqs, 439378781 residues

Total number of hits satisfying chosen parameters: 2443163

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : A. Geneseq 21.*

1: geneseqp1980s.*

2: geneseqp1990s.*

3: geneseqp2000s.*

4: geneseqp2001s.*

5: geneseqp2002s.*

6: geneseqp2003as.*

7: geneseqp2003bs.*

8: geneseqp2004s.*

9: geneseqp2005s.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	94	100.0	20	7	ADD93621 Streptococcus
2	94	100.0	20	9	ADK37244 Streptococcus
3	94	100.0	431	7	ADD93649 Streptococcus
4	94	100.0	431	7	ADD93650 Streptococcus
5	94	100.0	431	7	ADD93653 Streptococcus
6	94	100.0	431	9	ADK37272 Streptococcus
7	94	100.0	431	9	ADK37273 Streptococcus
8	94	100.0	431	9	ADK37276 Streptococcus
9	94	100.0	431	9	AE91500 Microbial
10	94	100.0	432	7	ADD93651 Streptococcus
11	94	100.0	432	7	ADD93652 Streptococcus
12	94	100.0	432	9	ADK37274 Streptococcus
13	94	100.0	432	9	ADK37275 Streptococcus
14	75	79.8	20	7	ADD93631 Streptococcus
15	75	79.8	20	9	ADK37254 Streptococcus
16	69	73.4	398	5	ABP25919 Streptococcus
17	69	73.4	398	8	ADR83884 S. pyogen
18	69	73.4	398	9	AE91542 Microbial
19	63	67.0	210	2	AA122579 Bacterial
20	63	67.0	392	6	ABU02747 S. pneumo
21	63	67.0	392	8	ADK47859 Streptococcus
22	63	67.0	392	8	ADT50227 S_pneumon
23	63	67.0	392	8	ADT50226 S_pneumon
24	63	67.0	392	8	ADT50165 S_pneumon

ALIGNMENTS

RESULT 1

ADD93621

ID ADD93621 standard; peptide; 20 AA.

XX AC ADD93621;

XX DT 29-JAN-2004 (first entry)

XX DE Streptococcus mutans glucan binding protein-B SVI peptide.

XX KW Glucan binding protein-B; GbpB; vaccine; anticaries; epitope; immunogen.

XX OS Streptococcus mutans.

XX PN WO2003075845-A2.

XX PD 18-SEP-2003.

XX PF 07-MAR-2003; 2003WO-US006962.

XX PR 07-MAR-2002; 2002US-0363209P.

XX PR 08-AUG-2002; 2002US-0402483P.

XX (FORS-) FORSYTH INST.

XX Smith DJ, Taubman MA;

XX WPI; 2003-845091/78.

Composition useful as vaccines for dental caries comprises a fragment of a glucan binding protein-B binding to a major histocompatibility complex class II protein.

Claim 12; Page 10; 49pp; English.

The present sequence is that of SYI peptide comprising amino acid residues 113-132 of the glucan binding protein-B (GbpB) of Streptococcus mutans strain SK32 ADD93649. The peptide binds to a major histocompatibility complex (MHC) class II protein. It was identified as a potential B cell epitope using a matrix-based algorithm for epitope prediction, which was used to search the primary amino acid sequence of GbpB for known MHC class II binding motifs. The peptide can be used in immunogenic compositions and subunit vaccines for dental caries. These comprise MHC class II protein-binding GbpB peptides covalently linked with peptide subunits (preferably from the catalytic domain) of a glucosyltransferase. The compositions are used in a claimed method of eliciting production of an antibody in a mammal. Dieptopic or

CC multiepitopic polypeptides can be prepared synthetically or by
 CC recombinant DNA technology. Antibodies raised against MHC class II
 CC binding fragments of GbpB can be used in passive immunisation.

XX
 SQ Sequence 20 AA;

Query Match 100.0%; Score 94; DB 7; Length 20;
 Best Local Similarity 100.0%; Pred. No. 1.9e-08;
 Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 KSNAAATSYINAIINSKVSVD 20

Db 1 KSNAAATSYINAIINSKVSVD 20

RESULT 2

ADX37244
 ID ADX37244 standard; peptide; 20 AA.

XX
 AC ADX37244;

DT 21-APR-2005 (first entry)

DE Streptococcus mutant glucan binding protein B peptide #1.

KW immunogenicity; immune stimulation; glucan binding protein-B;
 KW microparticle; major histocompatibility complex; tooth disease.

OS Streptococcus mutans.

XX US2005031633-A1.

PN 10-FEB-2005.

XX 09-MAR-2004; 2004US-00797821.

PR 13-APR-1998; 98US-0081550P.

PR 08-JAN-1999; 99US-0115142P.

PR 12-APR-1999; 99US-00290049.

PR 07-MAR-2002; 2002US-0363209P.

PR 08-AUG-2002; 2002US-0402483P.

PR 07-MAR-2003; 2003US-00383930.

XX (SMIT/) SMITH D J.

PA (TAUB/) TAUBMAN M A.

XX Smith DJ, Taubman MA;

PI WPI; 2005-151644/16.

DR

XX New composition comprising a fragment of a glucan binding protein-B
 PT (GbpB) that binds to MHC class II protein, and a biocompatible
 PT microparticle, useful for producing an antibody (claimed) for immunizing
 PT mammals against dental caries.

PS Claim 4; SEQ ID NO 1; 73pp; English.

XX The invention relates to a composition comprising a fragment of a glucan
 CC binding protein-B (GbpB) and a biocompatible microparticle, where the
 CC fragment binds to a major histocompatibility complex (MHC) class II
 CC protein. The composition is useful for producing an antibody for
 CC immunizing mammals against dental caries. This sequence corresponds to a
 CC Streptococcus mutans GbpB-derived peptide of the invention.

XX Sequence 20 AA;

Query Match 100.0%; Score 94; DB 9; Length 20;
 Best Local Similarity 100.0%; Pred. No. 1.9e-08;
 Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 KSNAAATSYINAIINSKVSVD 20

Db 1 KSNAAATSYINAIINSKVSVD 20

RESULT 3
 ADD93649
 ID ADD93649 standard; protein; 431 AA.
 XX
 AC ADD93649;
 XX
 DT 29-JAN-2004 (first entry)
 XX
 DE Streptococcus mutans glucan binding protein-B.
 XX
 KW Glucan binding protein-B; GbpB; vaccine; anticaries; epitope; immunogen.
 XX
 OS Streptococcus mutans.
 XX
 FH Key Location/Qualifiers
 FT Region 6..25
 FT /note= "HLA-binding peptide"
 FT Region 16..35
 FT /note= "HLA-binding peptide"
 FT Region 33..52
 FT /note= "HLA-binding peptide"
 FT Region 37..56
 FT /note= "HLA-binding peptide"
 FT Region 48..67
 FT /note= "HLA-binding peptide"
 FT Region 52..71
 FT /note= "HLA-binding peptide"
 FT Region 88..107
 FT /note= "HLA-binding peptide"
 FT Region 113..132
 FT /note= "HLA-binding peptide"
 FT Region 117..136
 FT /note= "HLA-binding peptide"
 FT Region 137..156
 FT /note= "HLA-binding peptide"
 FT Region 174..193
 FT /note= "HLA-binding peptide"
 FT Region 194..213
 FT /note= "HLA-binding peptide"
 FT Region 214..233
 FT /note= "HLA-binding peptide"
 FT Region 248..267
 FT /note= "HLA-binding peptide"
 FT Region 289..308
 FT /note= "HLA-binding peptide"
 FT Region 306..325
 FT /note= "HLA-binding peptide"
 FT Region 311..330
 FT /note= "HLA-binding peptide"
 FT Region 349..368
 FT /note= "HLA-binding peptide"
 FT Region 365..384
 FT /note= "HLA-binding peptide"
 FT Region 383..402
 FT /note= "HLA-binding peptide"
 FT Region 403..422
 FT /note= "HLA-binding peptide"

WO2003075845-A2.

XX 18-SEP-2003.

XX 07-MAR-2003; 2003WO-US006962.

PR 07-MAR-2002; 2002US-0363209P.

PR 08-AUG-2002; 2002US-0402483P.

PA (FORS-) FORSYTH INST.

XX Smith DJ, Taubman MA;

XX

DR WPI: 2003-845091/78.
 DR GENBANK; AY046410.
 CC Composition useful as vaccines for dental caries comprises a fragment of
 PT a glucan binding protein-B binding to a major histocompatibility complex
 PT class II protein.
 XX
 XX
 PS Claim 6; Page 7; 49pp; English.
 CC The present sequence is the protein sequence of the glucan binding
 CC protein-B (GbpB) of Streptococcus mutans strain SK32. The sequence
 CC includes a number of human leucocyte antigen (HLA)-binding epitopes. The
 CC invention provides immunogenic compositions and vaccines for dental
 CC caries. The compositions comprise major histocompatibility complex (MHC)
 CC class II protein-binding GbpB peptides, especially HLA-binding peptides,
 CC covalently linked with peptide subunits of a glucosyltransferase. The
 CC compositions are used in a claimed method of eliciting production of an
 CC antibody in a mammal. Diepitopic or multiepitopic polypeptides can be
 CC prepared synthetically or by recombinant DNA technology. Antibodies
 CC raised against MHC class II binding fragments of GbpB can be used in
 CC passive immunisation.
 XX
 SQ Sequence 431 AA;
 Query Match 100.0%; Score 94; DB 7; Length 431;
 Best Local Similarity 100.0%; Pred. No. 6.8e-07;
 Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 KSNATSYINAIINSKVSVD 20
 Db 113 KSNATSYINAIINSKVSVD 132
 RESULT 4
 ADD93650
 ID ADD93650 standard; protein; 431 AA.
 XX
 AC ADD93650;
 XX
 DT 29-JAN-2004 (first entry)
 XX
 DE Streptococcus mutans glucan binding protein-B.
 XX
 KW Glucan binding protein-B; GbpB; vaccine; anticaries; epitope; immunogen.
 XX
 OS Streptococcus mutans.
 XX
 PN WO2003075845-A2.
 XX
 PD 18-SEP-2003.
 XX
 PF 07-MAR-2003; 2003WO-US006962.
 XX
 PR 07-MAR-2002; 2002US-0363209P.
 PR 08-AUG-2002; 2002US-0402483P.
 XX
 PA (FORS-) FORSYTH INST.
 XX
 PI Smith DJ, Taubman MA;
 XX
 DR WPI: 2003-845091/78.
 DR GENBANK; AY046411.
 CC Composition useful as vaccines for dental caries comprises a fragment of
 PT a glucan binding protein-B binding to a major histocompatibility complex
 PT class II protein.
 XX
 XX
 PS Claim 5; Page 8; 49pp; English.
 CC The present sequence is the protein sequence of the glucan binding
 CC protein-B (GbpB) of Streptococcus mutans strain 3VF4. The sequence
 CC includes a number of human leucocyte antigen (HLA)-binding epitopes. The
 CC invention provides immunogenic compositions and vaccines for dental
 CC caries. The compositions comprise major histocompatibility complex (MHC)
 CC class II protein-binding GbpB peptides, especially HLA-binding peptides,
 CC covalently linked with peptide subunits of a glucosyltransferase. The
 CC compositions are used in a claimed method of eliciting production of an
 CC antibody in a mammal. Diepitopic or multiepitopic polypeptides can be
 CC prepared synthetically or by recombinant DNA technology. Antibodies
 CC raised against MHC class II binding fragments of GbpB can be used in
 CC passive immunisation.
 XX
 SQ Sequence 431 AA;
 Query Match 100.0%; Score 94; DB 7; Length 431;
 Best Local Similarity 100.0%; Pred. No. 6.8e-07;
 Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 KSNATSYINAIINSKVSVD 20
 Db 113 KSNATSYINAIINSKVSVD 132
 RESULT 5
 ADD93653
 ID ADD93653 standard; protein; 431 AA.
 XX
 AC ADD93653;
 XX
 DT 29-JAN-2004 (first entry)
 XX
 DE Streptococcus mutans glucan binding protein-B.
 XX
 KW Glucan binding protein-B; GbpB; vaccine; anticaries; epitope; immunogen.
 XX
 OS Streptococcus mutans.
 XX
 PN WO2003075845-A2.
 XX
 PD 18-SEP-2003.
 XX
 PF 07-MAR-2003; 2003WO-US006962.
 XX
 PR 07-MAR-2002; 2002US-0363209P.
 PR 08-AUG-2002; 2002US-0402483P.
 XX
 PA (FORS-) FORSYTH INST.
 XX
 PI Smith DJ, Taubman MA;
 XX
 DR WPI: 2003-845091/78.
 DR GENBANK; AY046414.
 CC Composition useful as vaccines for dental caries comprises a fragment of
 PT a glucan binding protein-B binding to a major histocompatibility complex
 PT class II protein.
 XX
 XX
 PS Claim 5; Page 8-9; 49pp; English.
 CC The present sequence is the protein sequence of the glucan binding
 CC protein-B (GbpB) of Streptococcus mutans strain 5SM3. The sequence
 CC includes a number of human leucocyte antigen (HLA)-binding epitopes. The
 CC invention provides immunogenic compositions and vaccines for dental
 CC caries. The compositions comprise major histocompatibility complex (MHC)
 CC class II protein-binding GbpB peptides, especially HLA-binding peptides,
 CC covalently linked with peptide subunits of a glucosyltransferase. The
 CC compositions are used in a claimed method of eliciting production of an
 CC antibody in a mammal. Diepitopic or multiepitopic polypeptides can be
 CC prepared synthetically or by recombinant DNA technology. Antibodies
 CC raised against MHC class II binding fragments of GbpB can be used in
 CC passive immunisation.
 XX
 SQ Sequence 431 AA;
 Query Match 100.0%; Score 94; DB 7; Length 431;
 Best Local Similarity 100.0%; Pred. No. 6.8e-07;

Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KSNAAATSYINAIINSKVSVD 20
 DB 113 KSNAAATSYINAIINSKVSVD 132

RESULT 6
 ADX37272
 ID ADX37272 standard; protein; 431 AA.
 XX AC ADX37272;
 XX DT 21-APR-2005 (first entry)
 XX DE Streptococcus mutant glucan binding protein B variant #1.
 XX KW immunogenicity; immune stimulation; glucan binding protein-B;
 KW microparticle; major histocompatibility complex; tooth disease.
 XX OS Streptococcus mutans.
 XX PN US2005031633-A1.
 XX PD 10-FEB-2005.
 XX PF 09-MAR-2004; 2004US-00797821.
 XX PR 13-APR-1998; 98US-0081550P.
 PR 08-JAN-1999; 99US-0115142P.
 PR 12-APR-1999; 99US-00290049.
 PR 07-MAR-2002; 2002US-0363209P.
 PR 08-AUG-2002; 2002US-0402483P.
 PR 07-MAR-2003; 2003US-00383930.
 XX (SMIT/) SMITH D J.
 PA (TAUB/) TAUBMAN M A.
 XX PI Smith DJ, Taubman MA;
 XX WPI; 2005-151644/16.
 XX New composition comprising a fragment of a glucan binding protein-B (GbpB) that binds to MHC class II protein, and a biocompatible microparticle, useful for producing an antibody (claimed) for immunizing mammals against dental caries.
 XX Claim 3; SEQ ID NO 29; 73pp; English.
 XX The invention relates to a composition comprising a fragment of a glucan binding protein-B (GbpB) and a biocompatible microparticle, where the fragment binds to a major histocompatibility complex (MHC) class II protein. The composition is useful for producing an antibody for immunizing mammals against dental caries. This sequence corresponds to a Streptococcus mutans GbpB protein of the invention.

Query Match 100.0%; Score 94; DB 9; Length 431;
 Best Local Similarity 100.0%; Pred. No. 6.8e-07;
 Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KSNAAATSYINAIINSKVSVD 20
 DB 113 KSNAAATSYINAIINSKVSVD 132

RESULT 7
 ADX37273
 ID ADX37273 standard; protein; 431 AA.
 XX AC ADX37273;
 XX DT 21-APR-2005 (first entry)
 XX DE Streptococcus mutant glucan binding protein B variant #5.
 XX KW immunogenicity; immune stimulation; glucan binding protein-B;
 KW microparticle; major histocompatibility complex; tooth disease.
 XX OS Streptococcus mutans.
 XX PN US2005031633-A1.
 XX PD 10-FEB-2005.

DT 21-APR-2005 (first entry)
 XX Streptococcus mutant glucan binding protein B variant #2.
 DE immunogenicity; immune stimulation; glucan binding protein-B;
 KW microparticle; major histocompatibility complex; tooth disease.
 XX OS Streptococcus mutans.
 XX PN US2005031633-A1.
 XX PD 10-FEB-2005.
 XX PF 09-MAR-2004; 2004US-00797821.
 XX PR 13-APR-1998; 98US-0081550P.
 PR 08-JAN-1999; 99US-0115142P.
 PR 12-APR-1999; 99US-00290049.
 PR 07-MAR-2002; 2002US-0363209P.
 PR 08-AUG-2002; 2002US-0402483P.
 PR 07-MAR-2003; 2003US-00383930.
 XX (SMIT/) SMITH D J.
 PA (TAUB/) TAUBMAN M A.
 XX PI Smith DJ, Taubman MA;
 XX WPI; 2005-151644/16.
 XX New composition comprising a fragment of a glucan binding protein-B (GbpB) that binds to MHC class II protein, and a biocompatible microparticle, useful for producing an antibody (claimed) for immunizing mammals against dental caries.
 XX Claim 3; SEQ ID NO 30; 73pp; English.
 XX The invention relates to a composition comprising a fragment of a glucan binding protein-B (GbpB) and a biocompatible microparticle, where the fragment binds to a major histocompatibility complex (MHC) class II protein. The composition is useful for producing an antibody for immunizing mammals against dental caries. This sequence corresponds to a Streptococcus mutans GbpB protein of the invention.

Query Match 100.0%; Score 94; DB 9; Length 431;
 Best Local Similarity 100.0%; Pred. No. 6.8e-07;
 Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KSNAAATSYINAIINSKVSVD 20
 DB 113 KSNAAATSYINAIINSKVSVD 132

RESULT 8
 ADX37276
 ID ADX37276 standard; protein; 431 AA.
 XX AC ADX37276;
 XX DT 21-APR-2005 (first entry)
 XX DE Streptococcus mutant glucan binding protein B variant #5.
 XX KW immunogenicity; immune stimulation; glucan binding protein-B;
 KW microparticle; major histocompatibility complex; tooth disease.
 XX OS Streptococcus mutans.
 XX PN US2005031633-A1.
 XX PD 10-FEB-2005.

PF 09-MAR-2004; 2004US-00797821.
 XX 13-APR-1998; 98US-0081550P.
 PR 08-JAN-1999; 99US-0115142P.
 PR 12-APR-1999; 99US-00290049.
 PR 07-MAR-2002; 2002US-0363209P.
 PR 08-AUG-2002; 2002US-0402483P.
 PR 07-MAR-2003; 2003US-00383930.
 XX (SMIT/) SMITH D J.
 PA (TAUB/) TAUBMAN M A.
 XX Smith DJ, Taubman MA;
 XX WPI; 2005-151644/16.
 DR New composition comprising a fragment of a glucan binding protein-B
 PT (GbpB) that binds to MHC class II protein, and a biocompatible
 PT microparticle, useful for producing an antibody (claimed) for immunizing
 PT mammals against dental caries.
 XX Claim 3; SEQ ID NO 33; 73pp; English.
 XX The invention relates to a composition comprising a fragment of a glucan
 CC binding protein-B (GbpB) and a biocompatible microparticle, where the
 CC fragment binds to a major histocompatibility complex (MHC) class II
 CC protein. The composition is useful for producing an antibody for
 CC immunizing mammals against dental caries. This sequence corresponds to a
 CC Streptococcus mutans GbpB protein of the invention.
 XX Sequence 431 AA;
 SQ Query Match 100.0%; Score 94; DB 9; Length 431;
 Best Local Similarity 100.0%; Pred. No. 6.8e-07;
 Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 KSNATSYINAINIINSKVSVD 20
 DB ||||||||||||||||||||
 113 KSNATSYINAINIINSKVSVD 132
 RESULT 9
 AEB91500
 ID AEB91500 standard; protein; 431 AA.
 XX AEB91500;
 XX 20-OCT-2005 (first entry)
 XX Microbial pathogen adhesin protein sequence, SEQ ID NO:210.
 DE algorithm; adhesin; pharmaceutical; vaccine; drug screening;
 KW bordetella pertussis infection; antibacterial; pneumonia;
 KW antiinflammatory; respiratory-gen.; gastric ulcer; antiulcer;
 KW gastrointestinal-gen.; urinary tract infection; antimicrobial; uropathic.
 XX Streptococcus mutans.
 OS WO2005076010-A2.
 PN 18-AUG-2005.
 XX 07-FEB-2005; 2005WO-IN000037.
 XX 06-FEB-2004; 2004IN-DE000173.
 PR 20-JUL-2004; 2004US-0589227P.
 XX (COUL) COUNCIL SCI & IND RES SOUTH AFRICA.
 XX Sachdeva G, Kumar K, Jain P, Brahmachari SK, Ramachandran S;
 PI WPI; 2005-597835/61.
 XX

PT Computational method for identifying adhesin and adhesin like molecules,
 PT comprises computing sequence-based attributes of protein sequences using
 PT neural network software and training an artificial neural network.
 XX Claim 16; SEQ ID NO 210; 402pp; English.
 XX The present invention relates to a computational method (M1) for
 CC identifying adhesin and adhesin-like proteins, by computing the sequence-
 CC based attributes of protein sequences using five attribute modules of a
 CC neural network software, training an artificial neural network (ANN) for
 CC each of the computed five attributes, and identifying the adhesin and
 CC adhesin-like proteins having probability of being an adhesin (pad) as
 CC equal or greater than 0.51. Also claimed is a set of 274 annotated genes
 CC encoding adhesin and adhesin-like proteins, having 274 fully defined 162-
 CC 11151 base pair (SEQ ID NO: 385-658) sequences; a set of 105 hypothetical
 CC genes encoding adhesin and adhesin-like proteins, having 105 fully
 CC defined 306-15676 base pairs (SEQ ID NO: 659-763); a set of 279 annotated
 CC adhesin and adhesin-like proteins, having 279 fully defined 53-3716 base
 CC pair (SEQ ID NO: 1-279); a set of 105 hypothetical adhesin and adhesin-
 CC like proteins, having 105 fully defined 106-5291 base pair (SEQ ID NO:
 CC 280-384) sequences; and a fully connected multilayer feed forward ANN (I)
 CC based on (M1). (M1) is useful for identifying adhesin and adhesin-like
 CC proteins, of therapeutic potential, and identifying and short-listing
 CC to eliminate diseases caused by various pathogenic organisms. (M1) is
 CC useful for identifying putative adhesins that are important in drug
 CC discovery and preventing therapeutics for whooping cough, pneumonia,
 CC gastric ulcer and urinary tract infections. (M1) identifies adhesins from
 CC distantly related organisms, and from bacteria belonging to a wide
 CC phylogenetic spectrum. (M1) is capable of predicting adhesive nature of
 CC unique proteins. The present sequence is a microbial pathogen adhesin
 CC protein sequence.
 XX Sequence 431 AA;
 SQ Query Match 100.0%; Score 94; DB 9; Length 431;
 Best Local Similarity 100.0%; Pred. No. 6.8e-07;
 Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 KSNATSYINAINIINSKVSVD 20
 DB ||||||||||||||||||||
 113 KSNATSYINAINIINSKVSVD 132
 RESULT 10
 ADD93651
 ID ADD93651 standard; protein; 432 AA.
 XX ADD93651;
 XX 29-JAN-2004 (first entry)
 XX Streptococcus mutans glucan binding protein-B.
 DE Glucan binding protein-B; GbpB; vaccine; anticaries; epitope; immunogen.
 XX Streptococcus mutans.
 OS WO2003075845-A2.
 PN 18-SEP-2003.
 XX 07-MAR-2003; 2003WO-US006962.
 XX 07-MAR-2002; 2002US-0363209P.
 PR 08-AUG-2002; 2002US-0402483P.
 XX (FORS-) FORSYTH INST.
 XX Smith DJ, Taubman MA;
 XX WPI; 2003-845091/78.
 DR GENBANK; AY046412.

XX Composition useful as vaccines for dental caries comprises a fragment of
 PT a glucan binding protein-B binding to a major histocompatibility complex
 PT class II protein.

PS Claim 5; Page 8; 49pp; English.

XX The present sequence is the protein sequence of the glucan binding
 CC protein-B (GbpB) of Streptococcus mutans strain 15JF2. The sequence
 CC includes a number of human leucocyte antigen (HLA)-binding epitopes. The
 CC invention provides immunogenic compositions and vaccines for dental
 CC caries. The compositions comprise major histocompatibility complex (MHC)
 CC class II protein-binding GbpB peptides, especially HLA-binding peptides,
 CC covalently linked with peptide subunits of a glucosyltransferase. The
 CC compositions are used in a claimed method of eliciting production of an
 CC antibody in a mammal. Diepitopic or multiepitopic polypeptides can be
 CC prepared synthetically or by recombinant DNA technology. Antibodies
 CC raised against MHC class II binding fragments of GbpB can be used in
 CC passive immunisation.

XX SQ Sequence 432 AA;

Query Match 100.0%; Score 94; DB 7; Length 432;
 Best Local Similarity 100.0%; Pred. No. 6.8e-07;
 Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KSNAAATSYINALINSKVSVD 20
 Db 113 KSNAAATSYINALINSKVSVD 132
 |||||

RESULT 11
 ADD93652
 ID ADD93652 standard; protein; 432 AA.

XX AC ADD93652;

XX DT 29-JAN-2004 (first entry)

XX DE Streptococcus mutans glucan binding protein-B.

XX KW Glucan binding protein-B; GbpB; vaccine; anticaries; epitope; immunogen.

XX OS Streptococcus mutans.

XX PN WO2003075845-A2.

XX PD 18-SEP-2003.

XX PF 07-MAR-2003; 2003WO-US0005962.

XX PR 07-MAR-2002; 2002US-0363209P.

XX PR 08-AUG-2002; 2002US-0402483P.

XX PA (FORS-) FORSYTH INST.

XX PI Smith DJ, Taubman MA;

XX PS WPI; 2003-845091/78.

XX DR GENBANK; AY046413.

XX Composition useful as vaccines for dental caries comprises a fragment of
 PT a glucan binding protein-B binding to a major histocompatibility complex
 PT class II protein.

PS Claim 5; Page 8; 49pp; English.

XX The present sequence is the protein sequence of the glucan binding
 CC protein-B (GbpB) of Streptococcus mutans strain 3SN1. The sequence
 CC includes a number of human leucocyte antigen (HLA)-binding epitopes. The
 CC invention provides immunogenic compositions and vaccines for dental
 CC caries. The compositions comprise major histocompatibility complex (MHC)
 CC class II protein-binding GbpB peptides, especially HLA-binding peptides,

CC covalently linked with peptide subunits of a glucosyltransferase. The
 CC compositions are used in a claimed method of eliciting production of an
 CC antibody in a mammal. Diepitopic or multiepitopic polypeptides can be
 CC prepared synthetically or by recombinant DNA technology. Antibodies
 CC raised against MHC class II binding fragments of GbpB can be used in
 CC passive immunisation.

XX SQ Sequence 432 AA;

Query Match 100.0%; Score 94; DB 7; Length 432;
 Best Local Similarity 100.0%; Pred. No. 6.8e-07;
 Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KSNAAATSYINALINSKVSVD 20
 Db 113 KSNAAATSYINALINSKVSVD 132
 |||||

RESULT 12
 ADX37274
 ID ADX37274 standard; protein; 432 AA.

XX AC ADX37274;

XX DT 21-APR-2005 (first entry)

XX DE Streptococcus mutant glucan binding protein B variant #3.

XX KW immunogenicity; immune stimulation; glucan binding protein-B;
 KW microparticle; major histocompatibility complex; tooth disease.

XX OS Streptococcus mutans.

XX PN US2005031633-A1.

XX PD 10-FEB-2005.

XX PF 09-MAR-2004; 2004US-00797821.

XX PR 13-APR-1998; 98US-0081550P.

XX PR 08-JAN-1999; 99US-0115142P.

XX PR 12-APR-1999; 99US-0029004P.

XX PR 07-MAR-2002; 2002US-0363209P.

XX PR 08-AUG-2002; 2002US-0402483P.

XX PR 07-MAR-2003; 2003US-00383930.

XX PA (SMIT/) SMITH D J.

XX PA (TAUB/) TAUBMAN M A.

XX PI Smith DJ, Taubman MA;

XX PS WPI; 2005-151644/16.

XX New composition comprising a fragment of a glucan binding protein-B
 PT (GbpB) that binds to MHC class II protein, and a biocompatible
 PT microparticle, useful for producing an antibody (claimed) for immunizing
 PT mammals against dental caries.

XX Claim 3; SEQ ID NO 31; 73pp; English.

XX The invention relates to a composition comprising a fragment of a glucan
 CC binding protein-B (GbpB) and a biocompatible microparticle, where the
 CC fragment binds to a major histocompatibility complex (MHC) class II
 CC protein. The composition is useful for producing an antibody for
 CC immunizing mammals against dental caries. This sequence corresponds to a
 CC Streptococcus mutans GbpB protein of the invention.

XX SQ Sequence 432 AA;

Query Match 100.0%; Score 94; DB 9; Length 432;
 Best Local Similarity 100.0%; Pred. No. 6.8e-07;
 Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

QY      1 KSNAAATSYINAIINSKVSVD 20
Db      |||||
      113 KSNAAATSYINAIINSKVSVD 132

RESULT 13
ADK37275
ID      ADK37275 standard; protein; 432 AA.
XX
AC      ADK37275;
XX
DT      21-APR-2005 (first entry)
XX
DE      Streptococcus mutant glucan binding protein B variant #4.
XX
KW      immunogenicity; immune stimulation; glucan binding protein-B;
KW      microparticle; major histocompatibility complex; tooth disease.
XX
OS      Streptococcus mutans.
XX
PN      US2005031633-A1.
XX
PD      10-FEB-2005.
XX
PF      09-MAR-2004; 2004US-00797821.
XX
PR      13-APR-1998; 98US-0081550P.
PR      08-JAN-1999; 99US-0115142P.
PR      12-APR-1999; 99US-00290049.
PR      07-MAR-2002; 2002US-0363209P.
PR      08-AUG-2002; 2002US-0402483P.
PR      07-MAR-2003; 2003US-00383930.
XX
PA      (SMIT/) SMITH D J.
PA      (TAUB/) TAUBMAN M A.
XX
PI      Smith DJ, Taubman MA;
XX
WPI; 2005-151644/16.
XX
DR      New composition comprising a fragment of a glucan binding protein-B
PT      (GbpB) that binds to MHC class II protein, and a biocompatible
PT      microparticle, useful for producing an antibody (claimed) for immunizing
PT      mammals against dental caries.
XX
PS      Claim 3; SEQ ID NO 32; 73pp; English.
XX
CC      The invention relates to a composition comprising a fragment of a glucan
CC      binding protein-B (GbpB) and a biocompatible microparticle, where the
CC      fragment binds to a major histocompatibility complex (MHC) class II
CC      protein. The composition is useful for producing an antibody for
CC      immunizing mammals against dental caries. This sequence corresponds to a
CC      Streptococcus mutans GbpB protein of the invention.
XX
SQ      Sequence 432 AA;

Query Match      100.0%; Score 94; DB 9; Length 432;
Best Local Similarity 100.0%; Pred. No. 6.8e-07;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 KSNAAATSYINAIINSKVSVD 20
Db      |||||
      113 KSNAAATSYINAIINSKVSVD 132

RESULT 14
ADD93631
ID      ADD93631 standard; peptide; 20 AA.
XX
AC      ADD93631;
XX
DT      29-JAN-2004 (first entry)
XX
DE      Streptococcus mutant glucan binding protein B peptide #11.
KW      immunogenicity; immune stimulation; glucan binding protein-B;
KW      microparticle; major histocompatibility complex; tooth disease.
OS      Streptococcus mutans.
XX
PN      US2005031633-A1.
XX
PD      10-FEB-2005.
XX
PF      07-MAR-2003; 2003WO-US006962.
PR      07-MAR-2002; 2002US-0363209P.
PR      08-AUG-2002; 2002US-0402483P.
XX
PA      (FORS-) FORSYTH INST.
XX
PI      Smith DJ, Taubman MA;
XX
WPI; 2003-845091/78.
XX
DR      Composition useful as vaccines for dental caries comprises a fragment of
PT      a glucan binding protein-B binding to a major histocompatibility complex
PT      class II protein.
XX
PS      Claim 11; Page 10; 49pp; English.
XX
CC      The present sequence is that of a peptide comprising amino acid residues
CC      117-136 of the glucan binding protein-B (GbpB) of Streptococcus mutans
CC      strain SK32 ADD93649. The peptide binds to a major histocompatibility
CC      complex (MHC) class II protein. It was identified as a potential B cell
CC      epitope using a matrix-based algorithm for epitope prediction, which was
CC      used to search the primary amino acid sequence of GbpB for known MHC
CC      class II binding motifs. The peptide can be used in immunogenic
CC      compositions and subunit vaccines for dental caries. These comprise MHC
CC      class II protein-binding GbpB peptides covalently linked with peptide
CC      subunits (preferably from the catalytic domain) of a glucosyltransferase.
CC      The compositions are used in a claimed method of eliciting production of
CC      an antibody in a mammal. Diepitopic or multipitopic polypeptides can be
CC      prepared synthetically or by recombinant DNA technology. Antibodies
CC      raised against MHC class II binding fragments of GbpB can be used in
CC      passive immunisation.
XX
SQ      Sequence 20 AA;

Query Match      79.8%; Score 75; DB 7; Length 20;
Best Local Similarity 100.0%; Pred. No. 2.8e-05;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      5 ATSYINAIINSKVSVD 20
Db      |||||
      1 ATSYINAIINSKVSVD 16

RESULT 15
ADK37254
ID      ADK37254 standard; peptide; 20 AA.
XX
AC      ADK37254;
XX
DT      21-APR-2005 (first entry)
XX
DE      Streptococcus mutant glucan binding protein B peptide #11.
KW      immunogenicity; immune stimulation; glucan binding protein-B;
KW      microparticle; major histocompatibility complex; tooth disease.
OS      Streptococcus mutans.
XX
PN      US2005031633-A1.
XX
PD      10-FEB-2005.
XX
PF      07-MAR-2003; 2003WO-US006962.
PR      07-MAR-2002; 2002US-0363209P.
PR      08-AUG-2002; 2002US-0402483P.
XX
PA      (FORS-) FORSYTH INST.
XX
PI      Smith DJ, Taubman MA;
XX
WPI; 2003-845091/78.
XX
DR      Composition useful as vaccines for dental caries comprises a fragment of
PT      a glucan binding protein-B binding to a major histocompatibility complex
PT      class II protein.
XX
PS      Claim 11; Page 10; 49pp; English.
XX
CC      The present sequence is that of a peptide comprising amino acid residues
CC      117-136 of the glucan binding protein-B (GbpB) of Streptococcus mutans
CC      strain SK32 ADD93649. The peptide binds to a major histocompatibility
CC      complex (MHC) class II protein. It was identified as a potential B cell
CC      epitope using a matrix-based algorithm for epitope prediction, which was
CC      used to search the primary amino acid sequence of GbpB for known MHC
CC      class II binding motifs. The peptide can be used in immunogenic
CC      compositions and subunit vaccines for dental caries. These comprise MHC
CC      class II protein-binding GbpB peptides covalently linked with peptide
CC      subunits (preferably from the catalytic domain) of a glucosyltransferase.
CC      The compositions are used in a claimed method of eliciting production of
CC      an antibody in a mammal. Diepitopic or multipitopic polypeptides can be
CC      prepared synthetically or by recombinant DNA technology. Antibodies
CC      raised against MHC class II binding fragments of GbpB can be used in
CC      passive immunisation.
XX
SQ      Sequence 20 AA;

```

PF 09-MAR-2004; 2004US-00797821.
 XX
 PR 13-APR-1998; 98US-0081550P.
 PR 08-JAN-1999; 99US-0115142P.
 PR 12-APR-1999; 99US-00290049.
 PR 07-MAR-2002; 2002US-0363209P.
 PR 08-AUG-2002; 2002US-0402483P.
 PR 07-MAR-2003; 2003US-00383930.
 XX
 PA (SMIT/) SMITH D J.
 PA (TAUB/) TAUBMAN M A.
 XX
 PI Smith DJ, Taubman MA;
 XX
 DR WPI; 2005-151644/16.
 XX
 PT New composition comprising a fragment of a glucan binding protein-B
 PT (GbpB) that binds to MHC class II protein, and a biocompatible
 PT microparticle, useful for producing an antibody (claimed) for immunizing
 PT mammals against dental caries.
 XX
 PS Claim 4; SEQ ID NO 11; 73pp; English.
 XX
 CC The invention relates to a composition comprising a fragment of a glucan
 CC binding protein-B (GbpB) and a biocompatible microparticle, where the
 CC fragment binds to a major histocompatibility complex (MHC) class II
 CC protein. The composition is useful for producing an antibody for
 CC immunizing mammals against dental caries. This sequence corresponds to a
 CC Streptococcus mutans GbpB-derived peptide of the invention.
 XX
 SQ Sequence 20 AA;

Query Match 79.8%; Score 75; DB 9; Length 20;
 Best Local Similarity 100.0%; Pred. No. 2.8e-05;
 Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 5 ATSYINAIINSKVSVD 20
 |||||
 Db 1 ATSYINAIINSKVSVD 16

Search completed: January 25, 2006, 19:01:39
 Job time : 70.2 secs

Result No.	Query			DB	ID	Description
	Score	Match	Length			
1	63	67.0	392	2	G95258	secreted 45 kd pro
2	63	67.0	392	2	B98124	general stress pro
3	59	62.8	456	2	E86903	hypothetical prote
4	59	62.8	461	2	JN0097	secreted 45K prote
5	48	51.1	507	2	S05542	hypothetical prote
6	46	48.9	351	2	B97273	uncharacterized pr
7	45	47.9	235	2	D69978	conserved hypothet
8	45	47.9	622	2	A61197	6-methylsalicylic
9	45	47.9	1774	2	S13178	6-methylsalicylic
10	44.5	47.3	500	2	E38338	methyl-accepting c
11	44	46.8	392	2	F96937	cell wall-associat
12	44	46.8	418	2	G90216	conserved hypothet
13	44	46.8	787	2	H71453	hypothetical prote
14	44	46.8	887	2	T52488	26S proteasome reg
15	43.5	46.3	1077	2	D90387	peptidase related
16	43	45.7	590	2	A41335	microbial metallopro
17	43	45.7	1309	2	T93379	sexual differentiation
18	42	44.7	145	2	A60038	vasoactive intesti
19	42	44.7	170	1	VRHU	vasoactive intesti
20	42	44.7	170	1	VRRT	vasoactive intesti
21	42	44.7	170	2	A60037	vasoactive intesti
22	42	44.7	274	2	B95972	probable sugar upt
23	42	44.7	659	2	F70175	rep helicase, sing
24	42	44.7	669	2	C96630	hypothetical prote
25	42	44.7	862	2	S53913	FKB2 protein - yea
26	41.5	44.1	203	2	S48995	hypothetical prote
27	41	43.6	73	2	D89782	hypothetical prote
28	41	43.6	228	2	AF0462	probable exported
29	41	43.6	374	2	T30429	late expression fa

C:Genetics:
A:Gene: SSO0686

Query Match 46.8%; Score 44; DB 2; Length 418;
Best Local Similarity 60.0%; Pred. No. 30;
Matches 9; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

QY 6 TSYINAINIINSKVSVD 20
||| |||:|:|:|
DB 206 TSYSELIINAQSVND 220

RESULT 13

H71453
Hypothetical protein PH0287 - Pyrococcus horikoshii
C:Species: Pyrococcus horikoshii
C:Date: 14-Aug-1998 #sequence_revision 14-Aug-1998 #text_change 12-Jul-2004
C:Accession: H71453
R:Kawarabayasi, Y.; Sawada, M.; Horikawa, H.; Haikawa, Y.; Hino, Y.; Yamamoto, S.; Sekin
M.; Ohfuku, Y.; Funahashi, I.; Tanaka, T.; Kudoh, Y.; Yamazaki, J.; Kushida, N.; Oguchi
DNA Res. 5, 55-76, 1998
A:Title: Complete sequence and gene organization of the genome of a hyper-thermophilic a
A:Reference number: A71000; PMID:98344137; PMID:9679194
A:Accession: H71453
A>Status: preliminary; nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-787 <RAW>
A:Cross-references: UNIPROT:O58025; UNIPARC:UPI0000062D86; GB:AP000001; NID:g3236128; PI
A:Experimental source: strain OT3
A>Note: this accession replaces an interim accession for a sequence replaced by GenBank
C:Genetics:
A:Gene: PH0287

Query Match 46.8%; Score 44; DB 2; Length 787;
Best Local Similarity 50.0%; Pred. No. 57;
Matches 8; Conservative 4; Mismatches 4; Indels 0; Gaps 0;

QY 5 ATSYINAINIINSKVSVD 20
| | | | | : | : | : |
DB 521 ADSYINNVFESESIAD 536

RESULT 14

T52488
26S proteasome regulatory particle chain 2 - fission yeast (Schizosaccharomyces pombe)
C:Species: Schizosaccharomyces pombe
C:Date: 24-Oct-2000 #sequence_revision 24-Oct-2000 #text_change 09-Jul-2004
C:Accession: T52488
R:Wilkinson, C.; Wallace, M.; Seeger, M.; Dubiel, W.; Gordon, C.
submitted to the EMBL Data Library, December 1996
A:Description: Characterization of Mts4 a subunit of the 26S protease demonstrates that
A:Reference number: Z26093
A:Accession: T52488
A>Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 1-887 <WIL>
A:Cross-references: UNIPROT:P87048; UNIPARC:UPI0000168FC6; EMBL:Y09819; PIDN:CAA70948.1
C:Genetics:
A:Gene: mts4
C:Superfamily: Caenorhabditis elegans hypothetical protein T22D1.9

Query Match 46.8%; Score 44; DB 2; Length 887;
Best Local Similarity 35.0%; Pred. No. 64;
Matches 7; Conservative 7; Mismatches 6; Indels 0; Gaps 0;

QY 1 KSNATSYINAINIINSKVSVD 20
| | | : | : | : | : | : |
DB 366 KQNLANTFVNALVAGYSND 385

RESULT 15

D90387
peptidase related protein [imported] - Sulfolobus solfataricus

C:Species: Sulfolobus solfataricus
C:Date: 24-May-2001 #sequence_revision 24-May-2001 #text_change 09-Jul-2004
C:Accession: D90387
R:She, Q.; Singh, R.K.; Confalonieri, P.; Zivanovic, Y.; Allard, G.; Awayez, M.J.; Chan-
Jong, I.; Jeffries, A.C.; Kozera, C.J.; Medina, N.; Peng, X.; Thi-Ngoc, H.P.; Redder, I.
arrett, R.A.; Ragan, M.A.; Sensen, C.W.; Van der Oost, J.
submitted to GenBank, April 2001
A:Description: Sulfolobus solfataricus complete genome.
A:Reference number: A99139
A:Accession: D90387
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-1077 <KUR>
A:Cross-references: UNIPROT:Q97WN1; UNIPARC:UPI000006467E; GB:AE006641; NID:gl3815479; I
C:Genetics:
A:Gene: SSO2181

Query Match 46.3%; Score 43.5; DB 2; Length 1077;
Best Local Similarity 55.0%; Pred. No. 94;
Matches 11; Conservative 5; Mismatches 3; Indels 1; Gaps 1;

QY 1 KSN-AATSYINAINIINSKVS 19
: | | : | | | : | : | : |
DB 86 QSNISATSYMNVILASGTVS 105

Search completed: January 25, 2006, 19:15:14
Job time : 10.1 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2006 Compugen Ltd.

OM protein - protein search, using sw model

Run on: January 25, 2006, 18:48:24 ; Search time 70.3 Seconds
(without alignments)
200.719 Million cell updates/sec

Title: US-10-797-821-1
Perfect score: 94
Sequence: 1 KSNATSYINAIINSKVSVD 20

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5
Searched: 2166443 seqs, 705528306 residues

Total number of hits satisfying chosen parameters: 2166443

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : UniProt_05.80:*
1: uniprot_sprot:*
2: uniprot_trembl:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	94	100.0	211	Q9ZAS7_STRMU	Q9ZAS7 streptococc
2	94	100.0	431	Q938V0_STRMU	Q938V0 streptococc
3	94	100.0	431	Q938V3_STRMU	Q938V3 streptococc
4	94	100.0	431	Q9AG98_STRMU	Q9AG98 streptococc
5	94	100.0	431	Q8DWM3_STRMU	Q8DWM3 streptococc
6	94	100.0	432	Q938V1_STRMU	Q938V1 streptococc
7	94	100.0	432	Q938V2_STRMU	Q938V2 streptococc
8	73	77.7	474	Q5M6K4_STRT2	Q5M6K4 streptococc
9	73	77.7	485	Q5M212_STRT1	Q5M212 streptococc
10	69	73.4	398	Q5XEL1_STRP6	Q5XEL1 streptococc
11	69	73.4	398	Q9A1Z8_STRP4	Q9A1Z8 streptococc
12	69	73.4	398	Q7CNQ7_STRP8	Q7CNQ7 streptococc
13	69	73.4	398	Q8P318_STRP3	Q8P318 streptococc
14	63	67.0	392	Q8DMY4_STRR6	Q8DMY4 streptococc
15	63	67.0	392	Q97N55_STRPN	Q97N55 streptococc
16	62	66.0	447	Q9AKA4_STRAG	Q9AKA4 streptococc
17	62	66.0	447	Q82H1_STRAS5	Q82H1 streptococc
18	62	66.0	447	Q8E7X9_STRAS3	Q8E7X9 streptococc
19	59	62.8	456	Q9CDJ1_LACIA	Q9CDJ1 lactococcu
20	59	62.8	461	1 USP45_ENTFC	P22865 lactococcu
21	52	55.3	524	Q9K2J9_ENTFC	Q9K2J9 enterococc
22	51	54.3	688	Q95UF4_ANCCA	Q95UF4 ancylobact
23	48	51.1	212	Q59ML3_CANAL	Q59ML3 candida alb
24	48	51.1	516	1 P54_ENTFC	P13692 enterococc
25	48	51.1	576	Q9KJ33_ENTHR	Q9KJ33 enterococc
26	48	51.1	1502	Q7R8L6_PLAYO	Q7R8L6 plasmodium
27	47	50.0	733	Q7RMP8_PLAYO	Q7RMP8 plasmodium
28	47	50.0	1038	Q5CJZ3_CRYHO	Q5CJZ3 cryptospori
29	47	50.0	1045	Q5CPW2_CRYPV	Q5CPW2 cryptospori
30	47	50.0	1435	Q54JF7_DICDI	Q54JF7 dictyosteli
31	47	50.0	1962	Q91QZ3_9VIRU	Q91QZ3 citrus leaf

32	46.5	49.5	825	2	Q6LW99_PROPR	Q6LW99 photobacter
33	46	48.9	351	2	Q97ES4_CLOAB	Q97ES4 clostridium
34	46	48.9	1000	2	Q7RQH5_PLAYO	Q7RQH5 plasmodium
35	46	48.9	1754	2	Q6MTF1_MYCMS	Q6MTF1 mycoplasma
36	46	48.9	1983	2	Q54GB1_DICDI	Q54GB1 dictyosteli
37	45	47.9	199	2	Q514G5_BYDVP	Q514G5 barley yell
38	45	47.9	199	2	Q9OCY8_9LUTE	Q9OCY8 barley yell
39	45	47.9	235	1	YRFD_BACSU	O05411 bacillus su
40	45	47.9	313	2	Q81HV3_BACCR	Q81HV3 bacillus ce
41	45	47.9	314	2	Q613A3_BACAN	Q613A3 bacillus an
42	45	47.9	314	2	Q4ML29_BACCE	Q4ML29 bacillus ce
43	45	47.9	314	2	Q6HND1_BACHK	Q6HND1 bacillus th
44	45	47.9	314	2	Q81V20_BACAN	Q81V20 bacillus an
45	45	47.9	314	2	Q63FW1_BAC CZ	Q63FW1 bacillus ce

ALIGNMENTS

RESULT 1
Q9ZAS7_STRMU Q9ZAS7_STRMU PRELIMINARY; PRT; 211 AA.
ID Q9ZAS7;
AC Q9ZAS7;
DT 01-MAY-1999 (Tremblrel. 10, Created)
DT 01-MAY-1999 (Tremblrel. 10, Last sequence update)
DT 01-OCT-2002 (Tremblrel. 22, Last annotation update)
DE Putative secreted protein (fragment).
OS Streptococcus mutans.
OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
OX Streptococcus.
NCBI_TaxID=1309;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=GS-5 Kuramitsu;
RA Peruzzi F., Piggot P.J., Daneo-Moore L.;
RL Submitted (NOV-1996) to the EMBL/GenBank/DBJ databases.
DR EMBL; U78607; AAD00288.1; -; Genomic_DNA.
FT NON_TER 211
SQ SEQUENCE 211 AA; 22803 MW; 4ACE331159CFAPC6 CRC64;

Query Match 100.0%; Score 94; DB 2; Length 211;
Best Local Similarity 100.0%; Pred. No. 2.4e-07;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KSNATSYINAIINSKVSVD 20
|||||
DB 113 KSNATSYINAIINSKVSVD 132

RESULT 2
Q938V0_STRMU Q938V0_STRMU PRELIMINARY; PRT; 431 AA.
ID Q938V0;
AC Q938V0;
DT 01-DEC-2001 (Tremblrel. 19, Created)
DT 01-DEC-2001 (Tremblrel. 19, Last sequence update)
DT 01-MAR-2004 (Tremblrel. 26, Last annotation update)
DE Glucan-binding protein B.
OS Streptococcus mutans.
OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
OX Streptococcus.
NCBI_TaxID=1309;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=5SM3;
RA Jin S., Duncan M.J., Taubman M.A., Smith D.J.;
RT "Cloning of the gbpB gene from Streptococcus mutans."; J. Dent. Res. 79:224-224 (2000).
RN [2]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=5SM3;
RX MEDLINE=21481971; PubMed=11598068;
DOI=10.1128/JAI.69.11.6931-6941.2001;
RA Mattos-Graner R.O., Jin S., King W.F., Chen T., Smith D.J.,

```

RA Duncan M.J.;
RT "Cloning of the Streptococcus mutans gene encoding glucan binding
RT protein B and analysis of genetic diversity and protein production in
RT clinical isolates.";
RL Infect. Immun. 69:6931-6941(2001).
DR EMBL; AY046414; AAK94504.1; -; Genomic_DNA.
DR InterPro; IPR007921; CHAP.
DR Pfam; PF05257; CHAP; 1.
DR PRINTS; PR01852; SIBAPROTEIN.
DR PROSITE; PS50911; CHAP; 1.
SQ SEQUENCE 431 AA; 44650 MW; 05D38D8B8C4609F CRC64;

Query Match 100.0%; Score 94; DB 2; Length 431;
Best Local Similarity 100.0%; Pred. No. 5.1e-07;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KSNAAATSYINAIINSKVSVD 20
Db 113 KSNAAATSYINAIINSKVSVD 132

RESULT 3
Q938V3_STRMU PRELIMINARY; PRT; 431 AA.
AC Q938V3;
DT 01-DEC-2001 (TrEMBLrel. 19, Created)
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Glucan-binding protein B.
OS Streptococcus mutans.
OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
OC Streptococcus.
OX * NCBI_TaxID=1309;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=GS-5;
RA Jin S., Duncan M.J., Taubman M.A., Smith D.J.;
RT "Cloning of the gbpB gene from Streptococcus mutans.";
RL J. Dent. Res. 79:224-224(2000).
RN [2]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=GS-5;
RX DOI=10.1128/IAI.69.11.6931-6941.2001;
RX DOI=10.1128/IAI.69.11.6931-6941.2001;
RA Mattos-Graner R.O., Jin S., King W.F., Chen T., Smith D.J.,
Duncan M.J.;
RT "Cloning of the Streptococcus mutans gene encoding glucan binding
RT protein B and analysis of genetic diversity and protein production in
RT clinical isolates.";
RL Infect. Immun. 69:6931-6941(2001).
DR EMBL; AY046410; AAK94500.1; -; Genomic_DNA.
DR InterPro; IPR007921; CHAP.
DR Pfam; PF05257; CHAP; 1.
DR PRINTS; PR01852; SIBAPROTEIN.
DR PROSITE; PS50911; CHAP; 1.
SQ SEQUENCE 431 AA; 44620 MW; 464FE3B563FB7B51 CRC64;

Query Match 100.0%; Score 94; DB 2; Length 431;
Best Local Similarity 100.0%; Pred. No. 5.1e-07;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KSNAAATSYINAIINSKVSVD 20
Db 113 KSNAAATSYINAIINSKVSVD 132

RESULT 4
Q9AG98_STRMU PRELIMINARY; PRT; 431 AA.
AC Q9AG98;
DT 01-JUN-2001 (TrEMBLrel. 17, Created)

```

```

DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
DT 01-FEB-2005 (TrEMBLrel. 29, Last annotation update)
DE Immunodominant glycoprotein IDG-60 (Glucan-binding protein B).
GN Name=saga;
OS Streptococcus mutans.
OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
OC Streptococcus.
OX NCBI_TaxID=1309;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=GS-5;
RX DOI=10.1128/IAI.69.4.2493-2501.2001;
RX DOI=10.1128/IAI.69.4.2493-2501.2001;
RA Chia J.-S., Lee Y.-Y., Huang P.T., Chen J.Y.;
RT "Identification of stress-responsive genes in Streptococcus mutans by
RT differential display reverse transcription-PCR.";
RL Infect. Immun. 69:2493-2501(2001).
RN [2]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=GS-5;
RX DOI=10.1128/IAI.69.11.6987-6998.2001;
RX DOI=10.1128/IAI.69.11.6987-6998.2001;
RA Chia J.-S., Chang L.-Y., Shun C.T., Chang Y.-Y., Chen J.-Y.;
RT "A 60-kilodalton immunodominant glycoprotein is essential for cell
RT wall integrity and the maintenance of cell shape in Streptococcus
RT mutans.";
RL Infect. Immun. 69:6987-6998(2001).
RN [3]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=GS-5;
RA Chia J.-S., Chang L.-Y., Lee Y.-Y., Chen J.-Y.;
RT "Cloning of the gbpB gene from Streptococcus mutans.";
RL J. Dent. Res. 79:224-224(2000).
RN [4]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=3VF4;
RA Jin S., Duncan M.J., Taubman M.A., Smith D.J.;
RT "Cloning of the gbpB gene from Streptococcus mutans.";
RN [5]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=3VF4;
RX DOI=10.1128/IAI.69.11.6931-6941.2001;
RX DOI=10.1128/IAI.69.11.6931-6941.2001;
RA Mattos-Graner R.O., Jin S., King W.F., Chen T., Smith D.J.,
Duncan M.J.;
RT "Cloning of the Streptococcus mutans gene encoding glucan binding
RT protein B and analysis of genetic diversity and protein production in
RT clinical isolates.";
RL Infect. Immun. 69:6931-6941(2001).
DR EMBL; AF338445; AAK08104.1; -; Genomic_DNA.
DR EMBL; AY046411; AAK94501.1; -; Genomic_DNA.
DR InterPro; IPR007921; CHAP.
DR Pfam; PF05257; CHAP; 1.
DR PRINTS; PR01852; SIBAPROTEIN.
DR PROSITE; PS50911; CHAP; 1.
SQ SEQUENCE 431 AA; 44592 MW; 3EBE21FC5B47232E CRC64;

Query Match 100.0%; Score 94; DB 2; Length 431;
Best Local Similarity 100.0%; Pred. No. 5.1e-07;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KSNAAATSYINAIINSKVSVD 20
Db 113 KSNAAATSYINAIINSKVSVD 132

RESULT 5
Q8DWM3_STRMU PRELIMINARY; PRT; 431 AA.
AC Q8DWM3;
DT 01-MAR-2003 (TrEMBLrel. 23, Created)
DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)

```

DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
 DE Putative secreted antigen GbpB/SagA; putative peptidoglycan
 DE hydrolase.
 GN Name=gbpB; OrderedLocusNames=SMU.22;
 OS Streptococcus mutans.
 OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
 OC Streptococcus.
 OX NCBI_TaxID=1309;
 RN [1]
 RP NUCLEOTIDE SEQUENCE.
 RC STRAIN=UAI59 / ATCC 700610 / Serotype C;
 RX MEDLINE=22295063; PubMed=12397186; DOI=10.1073/pnas.172501299;
 RA Ajdic D.J., McShane W.M., McLaughlin R.E., Savic G., Chang J.,
 RA Carson M.B., Primeaux C., Tian R., Kenton S., Jia H.G., Lin S.P.,
 RA Qian Y., Li S., Zhu H., Najjar F.Z., Lai H., White J., Roe B.A.,
 RA Ferretti J.J.;
 RT "Genome sequence of Streptococcus mutans UAI59, a cariogenic dental
 RT pathogen.";
 RL Proc. Natl. Acad. Sci. U.S.A. 99:14434-14439 (2002).
 DR EMBL; AB014855; AAN57811.1; -; Genomic_DNA.
 DR InterPro; IPR007921; CHAP.
 DR InterPro; IPR009148; Siba.
 DR Pfam; PF05257; CHAP; 1.
 DR PRINTS; PR01852; SIBAPROTEIN.
 DR PROSITE; PS50911; CHAP; 1.
 KW Complete proteome.
 SQ SEQUENCE 431 AA; 44620 MW; 2D1CA685248CCD3E CRC64;

Query Match 100.0%; Score 94; DB 2; Length 431;
 Best Local Similarity 100.0%; Pred. No. 5.1e-07;
 Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KSNAAATSYINAIINSKVSVD 20
 |||||
 Db 113 KSNAAATSYINAIINSKVSVD 132

RESULT 6
 Q938V1 STRMU
 ID Q938V1_STRMU PRELIMINARY; PRT; 432 AA.
 AC Q938V1;
 DT 01-DEC-2001 (TrEMBLrel. 19, Created)
 DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
 DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
 DE Glucan-binding protein B.
 OS Streptococcus mutans.
 OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
 OC Streptococcus.
 OX NCBI_TaxID=1309;
 RN [1]
 RP NUCLEOTIDE SEQUENCE.
 RC STRAIN=3SN1;
 RA Jin S., Duncan M.J., Taubman M.A., Smith D.J.;
 RT "Cloning of the gbpB gene from Streptococcus mutans.";
 RL J. Dent. Res. 79:224-224 (2000).
 RN [2]
 RP NUCLEOTIDE SEQUENCE.
 RC STRAIN=15JP2;
 RX MEDLINE=21481971; PubMed=11598068;
 DOI=10.1128/JAI.69.11.6931-6941.2001;
 RA Mattos-Graner R.O., Jin S., King W.F., Chen T., Smith D.J.,
 RA Duncan M.J.;
 RT "Cloning of the Streptococcus mutans gene encoding glucan binding
 RT protein B and analysis of genetic diversity and protein production in
 RT clinical isolates.";
 RL Infect. Immun. 69:6931-6941 (2001).
 DR EMBL; AY046412; AAK94502.1; -; Genomic_DNA.
 DR InterPro; IPR007921; CHAP.
 DR InterPro; IPR009148; Siba.
 DR Pfam; PF05257; CHAP; 1.
 DR PRINTS; PR01852; SIBAPROTEIN.
 DR PROSITE; PS50911; CHAP; 1.
 SQ SEQUENCE 432 AA; 44648 MW; E769B2504ABE50E9 CRC64;

Query Match 100.0%; Score 94; DB 2; Length 432;
 Best Local Similarity 100.0%; Pred. No. 5.2e-07;
 Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KSNAAATSYINAIINSKVSVD 20
 |||||
 Db 113 KSNAAATSYINAIINSKVSVD 132

RESULT 8
 Q5M6K4 STRT2
 ID Q5M6K4_STRT2 PRELIMINARY; PRT; 474 AA.
 AC Q5M6K4;
 DT 01-FEB-2005 (TrEMBLrel. 29, Created)
 DT 01-FEB-2005 (TrEMBLrel. 29, Last sequence update)
 DT 13-SEP-2005 (TrEMBLrel. 31, Last annotation update)
 DE Glucan binding protein (PcsB).
 GN Name=pcsB; OrderedLocusNames=stu0022;
 OS Streptococcus thermophilus (strain ATCC BAA-250 / LMG 18311).
 OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
 OC Streptococcus.
 OX NCBI_TaxID=264199;
 RN [1]
 RP NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
 RX PubMed=15543133; DOI=10.1038/nbt1034;
 RA Bolotin A., Quinquis B., Renaut P., Sorokin A., Ehrlich S.D.,

Query Match 100.0%; Score 94; DB 2; Length 432;
 Best Local Similarity 100.0%; Pred. No. 5.2e-07;
 Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KSNAAATSYINAIINSKVSVD 20
 |||||
 Db 113 KSNAAATSYINAIINSKVSVD 132

RESULT 7
 Q938V2 STRMU
 ID Q938V2_STRMU PRELIMINARY; PRT; 432 AA.
 AC Q938V2;
 DT 01-DEC-2001 (TrEMBLrel. 19, Created)
 DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
 DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
 DE Glucan-binding protein B.
 OS Streptococcus mutans.
 OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
 OC Streptococcus.
 OX NCBI_TaxID=1309;
 RN [1]
 RP NUCLEOTIDE SEQUENCE.
 RC STRAIN=15JP2;
 RA Jin S., Duncan M.J., Taubman M.A., Smith D.J.;
 RT "Cloning of the gbpB gene from Streptococcus mutans.";
 RL J. Dent. Res. 79:224-224 (2000).
 RN [2]
 RP NUCLEOTIDE SEQUENCE.
 RC STRAIN=15JP2;
 RX MEDLINE=21481971; PubMed=11598068;
 DOI=10.1128/JAI.69.11.6931-6941.2001;
 RA Mattos-Graner R.O., Jin S., King W.F., Chen T., Smith D.J.,
 RA Duncan M.J.;
 RT "Cloning of the Streptococcus mutans gene encoding glucan binding
 RT protein B and analysis of genetic diversity and protein production in
 RT clinical isolates.";
 RL Infect. Immun. 69:6931-6941 (2001).
 DR EMBL; AY046412; AAK94502.1; -; Genomic_DNA.
 DR InterPro; IPR007921; CHAP.
 DR InterPro; IPR009148; Siba.
 DR Pfam; PF05257; CHAP; 1.
 DR PRINTS; PR01852; SIBAPROTEIN.
 DR PROSITE; PS50911; CHAP; 1.
 SQ SEQUENCE 432 AA; 44648 MW; E769B2504ABE50E9 CRC64;

Query Match 100.0%; Score 94; DB 2; Length 432;
 Best Local Similarity 100.0%; Pred. No. 5.2e-07;
 Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KSNAAATSYINAIINSKVSVD 20
 |||||
 Db 113 KSNAAATSYINAIINSKVSVD 132

RESULT 8
 Q5M6K4 STRT2
 ID Q5M6K4_STRT2 PRELIMINARY; PRT; 474 AA.
 AC Q5M6K4;
 DT 01-FEB-2005 (TrEMBLrel. 29, Created)
 DT 01-FEB-2005 (TrEMBLrel. 29, Last sequence update)
 DT 13-SEP-2005 (TrEMBLrel. 31, Last annotation update)
 DE Glucan binding protein (PcsB).
 GN Name=pcsB; OrderedLocusNames=stu0022;
 OS Streptococcus thermophilus (strain ATCC BAA-250 / LMG 18311).
 OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
 OC Streptococcus.
 OX NCBI_TaxID=264199;
 RN [1]
 RP NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
 RX PubMed=15543133; DOI=10.1038/nbt1034;
 RA Bolotin A., Quinquis B., Renaut P., Sorokin A., Ehrlich S.D.,

Db 110 QNNTATSYINILNKSIS 129

RESULT 10
 Q5XEL1 STRP6
 ID Q5XEL1_STRP6 PRELIMINARY; PRT; 398 AA.
 AC Q5XEL1;
 DT 25-OCT-2004 (TREMBLrel. 28, Created)
 DT 25-OCT-2004 (TREMBLrel. 28, Last sequence update)
 DT 25-OCT-2004 (TREMBLrel. 28, Last annotation update)
 DE Secreted protein.
 GN OrderedLocusNames=M6_Spy0017;
 OS Streptococcus pyogenes (serotype M6).
 OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
 OC Streptococcus.
 OX NCBI_TaxID=301450;
 [1]
 RN NUCLEOTIDE SEQUENCE.
 RC STRAIN=MGAS10394;
 RK PubMed=15272401; DOI=10.1086/422697;
 RA Banks D.J., Delcelo S.P., Barbian K.D., Beres S.B., Phillips L.E.,
 RA Voyich J.M., Porcile F.R., Martin J.M., Somerville G.A., Musser J.M.;
 RT "Progress toward characterization of the group A Streptococcus
 RT metagenome: complete genome sequence of a macrolide-resistant serotype
 RT M6 strain."; Dis. 190:727-738(2004).
 RL J. Infect. Dis. 190:727-738(2004).
 DR EMBL; CP000003; AAT86152.1; -; Genomic_DNA.
 DR InterPro; IPR007921; CHAP.
 DR InterPro; IPR009148; Siba.
 DR Pfam; PF05257; CHAP; 1.
 DR PRINTS; PR01852; SIBAPROTEIN.
 DR PROSITE; PS0911; CHAP; 1.
 KW Complete proteome.
 SQ SEQUENCE 398 AA; 40208 MW; 5DCEDA78CB863B60 CRC64;

Query Match 73.4%; Score 69; DB 2; Length 398;
 Best Local Similarity 60.0%; Pred. No. 0.0076;
 Matches 12; Conservative 5; Mismatches 3; Indels 0; Gaps

Oy 1 KSNATSYINAIINSKVS 20
 | : | | | : | | | : | |
 110 KNNETSGYNALNKSISD 129

Db 110 QNNTATSYINILNKSIS 129

RESULT 11
 Q9A1Z8 STRP7
 ID Q9A1Z8_STRP7 PRELIMINARY; PRT; 398 AA.
 AC Q9A1Z8; Q7BH59;
 DT 01-JUN-2001 (TREMBLrel. 17, Created)
 DT 01-JUN-2001 (TREMBLrel. 17, Last sequence update)
 DT 10-MAY-2005 (TREMBLrel. 30, Last annotation update)
 DE Secreted protein Siba precursor.
 GN OrderedLocusNames=SPY0019;
 OS Streptococcus pyogenes.
 OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
 OC Streptococcus.
 OX NCBI_TaxID=1314;
 [1]
 RN NUCLEOTIDE SEQUENCE.
 RC STRAIN=SF370 / ATCC 700294 / Serotype M1;
 RX MEDLINE=21192694; PubMed=11296396; DOI=10.1073/pnas.071559398;
 RA Ferretti J.J., McGhan W.M., Ajdic D.J., Savic G., Lyon K.,
 RA Priemetti C., Sezate S., Suvoarov A.N., Kenton S., Lai H.S., Lin S.P.,
 RA Qian Y., Jia H.G., Najjar F.Z., Ren Q., Zhu H., Song L., White J.,
 RA Yuan X., Clifton S.W., Roe B.A., McLaughlin R.E.;
 RT "Complete genome sequence of an M1 strain of Streptococcus pyogenes.";
 RT Proc. Natl. Acad. Sci. U.S.A. 98:4658-4663(2001).
 RL [2]
 RN NUCLEOTIDE SEQUENCE.
 RA Pagan P.K., Reinscheid D., Gottschalk B., Chhatwal G.S.;
 RT "Identification and characterization of a novel secreted protein from
 RT group A streptococcus";
 RL Submitted (NOV-2000) to the EMBL/GenBank/DBJ databases

DR EMBL; AE006474; AAK33158.1; -; Genomic_DNA.
 DR EMBL; AF319999; AAL73135.1; -; Genomic_DNA.
 DR InterPro; IPR007921; CHAP.
 DR InterPro; IPR009148; Siba.
 DR Pfam; PF05257; CHAP; 1.
 DR PRINTS; PR01852; SIBAPROTEIN.
 DR PROSITE; PS050911; CHAP; 1.
 KW Complete proteome; Signal.
 FT SIGNAL 1 23 Potential.
 SQ SEQUENCE 398 AA; 41899 MW; 28A9B3F7195E969B CRC64;

Query Match 73.4%; Score 69; DB 2; Length 398;
 Best Local Similarity 60.0%; Pred. No. 0.0076;
 Matches 12; Conservative 5; Mismatches 3; Indels 0; Gaps 0;

OY 1 KSNATSYINAINLSKVSVD 20
 Db 110 KNETSGYINALLNSKSID 129
 :|: : ||||:||||:|

RESULT 12

ID Q7CNQ7 STRP8 PRELIMINARY; PRT; 398 AA.
 AC Q7CNQ7;
 DT 05-JUL-2004 (TrEMBLrel. 27, Created)
 DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
 DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
 DE Putative secreted protein.
 GN OrderedLocusNames=spvM18_0020;
 OS Streptococcus pyogenes (serotype M18).
 OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
 OC Streptococcus.
 OX NCBI_TaxID=301451;
 RN [1]
 RP NUCLEOTIDE SEQUENCE.

RC STRAIN=MGAS232 / Serotype M18;
 RX MEDLINE=21927593; PubMed=11917108; DOI=10.1073/pnas.062526099;
 RA Smoot J.C., Barbian K.D., Van Gompel J.J., Smoot L.M., Chaussee M.S.,
 RA Sylva G.L., Sturdevant D.B., Ricklefs S.M., Porcella S.F.,
 RA Parkins L.D., Beres S.B., Campbell D.S., Smith T.M., Zhang Q.,
 RA Kapur V., Daly J.A., Veasy L.G., Musser J.M.;
 RT "Genome sequence and comparative microarray analysis of serotype M18
 group A Streptococcus strains associated with acute rheumatic fever"
 RT outbreaks";
 RL Proc. Natl. Acad. Sci. U.S.A. 99:4668-4673 (2002).
 DR EMBL; AE009955; AAL96849.1; -; Genomic_DNA.
 DR InterPro; IPR007921; CHAP.
 DR InterPro; IPR009148; Siba.
 DR Pfam; PF05257; CHAP; 1.
 DR PRINTS; PR01852; SIBAPROTEIN.
 DR PROSITE; PS050911; CHAP; 1.
 KW Complete proteome.
 SQ SEQUENCE 398 AA; 42028 MW; 5DCEDA78CB863B60 CRC64;

Query Match 73.4%; Score 69; DB 2; Length 398;
 Best Local Similarity 60.0%; Pred. No. 0.0076;
 Matches 12; Conservative 5; Mismatches 3; Indels 0; Gaps 0;

OY 1 KSNATSYINAINLSKVSVD 20
 Db 110 KNETSGYINALLNSKSID 129
 :|: : ||||:||||:|

RESULT 13

ID Q8P318 STRP3 PRELIMINARY; PRT; 398 AA.
 AC Q8P318; Q7CFL7;
 DT 01-OCT-2002 (TrEMBLrel. 22, Created)
 DT 01-OCT-2002 (TrEMBLrel. 22, Last sequence update)
 DT 13-SEP-2005 (TrEMBLrel. 31, Last annotation update)
 DE Putative secreted protein.
 GN OrderedLocusNames=SPS0015, SpyM3_0014;
 OS Streptococcus pyogenes (serotype M3).

OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
 OC Streptococcus.
 OX NCBI_TaxID=301448;
 RN [1]
 RP NUCLEOTIDE SEQUENCE.

RC STRAIN=SSI-1 / Serotype M3;
 RX MEDLINE=22683278; PubMed=12799345; DOI=10.1101/gr.1096703;
 RA Nakagawa I., Kurokawa K., Yamashita A., Nakata M., Tomiyasu Y.,
 RA Okahashi N., Kawabata S., Yamazaki K., Shiba T., Yasunaga T.,
 RA Hayashi H., Hattori M., Hamada S.;
 RT "Genome sequence of an M3 strain of Streptococcus pyogenes reveals a
 large-scale genomic rearrangement in invasive strains and new insights
 into phase evolution.";
 RL Genome Res. 13:1042-1055 (2003).
 RN [2]

RP NUCLEOTIDE SEQUENCE.
 RC STRAIN=MGAS315 / Serotype M3;
 RX MEDLINE=22133808; PubMed=12122206; DOI=10.1073/pnas.152298499;
 RA Beres S.B., Sylva G.L., Barbian K.D., Lei B., Hoff J.S.,
 RA Mammarella N.D., Liu M.-Y., Smoot J.C., Porcella S.F., Parkins L.D.,
 RA Campbell D.S., Smith T.M., McCormick J.K., Leung D.Y.M.,
 RA Schlievert P.M., Musser J.M.;
 RT "Genome sequence of a serotype M3 strain of group A Streptococcus:
 phage-encoded toxins, the high-virulence phenotype, and clone
 emergence";
 RL Proc. Natl. Acad. Sci. U.S.A. 99:10078-10083 (2002).
 DR EMBL; BA000034; BAC63110.1; -; Genomic_DNA.
 DR EMBL; AE014136; AAM78621.1; -; Genomic_DNA.
 DR InterPro; IPR007921; CHAP.
 DR InterPro; IPR009148; Siba.
 DR Pfam; PF05257; CHAP; 1.
 DR PRINTS; PR01852; SIBAPROTEIN.
 DR PROSITE; PS050911; CHAP; 1.
 KW Complete proteome.
 SQ SEQUENCE 398 AA; 42028 MW; 5DCEDA78CB863B60 CRC64;

Query Match 73.4%; Score 69; DB 2; Length 398;
 Best Local Similarity 60.0%; Pred. No. 0.0076;
 Matches 12; Conservative 5; Mismatches 3; Indels 0; Gaps 0;

OY 1 KSNATSYINAINLSKVSVD 20
 Db 110 KNETSGYINALLNSKSID 129
 :|: : ||||:||||:|

RESULT 14

ID Q8DMY4 STRR6 PRELIMINARY; PRT; 392 AA.
 AC Q8DMY4;
 DT 01-MAR-2003 (TrEMBLrel. 23, Created)
 DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
 DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
 DE General stress protein GSP-781.
 GN Name=gsp-781; OrderedLocusNames=spr2021;
 OS Streptococcus pneumoniae (strain ATCC BAA-255 / R6).
 OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
 OC Streptococcus.
 OX NCBI_TaxID=171101;
 RN [1]
 RP NUCLEOTIDE SEQUENCE.

RX MEDLINE=21429245; PubMed=11544234;
 RX DOI=10.1128/JB.183.19.5709-5717.2001;
 RA Hoskins J., Alborn W.E. Jr., Arnold J., Blaszcak L.C., Burgett S.,
 RA DeHoff B.S., Betrem S.T., Fritz L., Fu D.-J., Fuller W., Geringer C.,
 RA Gilmore R., Glass J.S., Khoja H., Kraft A.R., Lagace R.E.,
 RA LeBlanc D.J., Lee L.N., Lefkowitz E.J., Lu J., Matsushima P.,
 RA McAhren S.M., McHenry M., McLeaster K., Mundy C.W., Nicass T.I.,
 RA Norris F.H., O'Garra M., Peery R.B., Robertson G.T., Rockey P.,
 RA Sun P.-M., Winkler M.E., Yang Y., Young-Bellido M., Zhao G.,
 RA Zook C.A., Baltz R.H., Jaskunas S.R., Rostek P.R. Jr., Skatrud P.L.,
 RA Glass J.I.;
 RT "Genome of the bacterium Streptococcus pneumoniae strain R6";
 RL J. Bacteriol. 183:5709-5717 (2001).

```

DR EMBL; AE008566; AAL00823.1; -; Genomic_DNA.
DR PIR; B98124; B98124.
DR PIR; G95258; G95258.
DR InterPro; IPR007921; CHAP.
DR InterPro; IPR009148; SIDA.
DR Pfam; PF05257; CHAP; 1.
DR PRINTS; PR01852; SIBAPROTEIN.
DR PROSITE; PS0911; CHAP; 1.
KW Complete proteome.
SQ SEQUENCE 392 AA; 41697 MW; B5A5A860FAEA77DD CRC64;

Query Match 67.0%; Score 63; DB 2; Length 392;
Best Local Similarity 57.9%; Pred. No. 0.077;
Matches 11; Conservative 5; Mismatches 3; Indels 0; Gaps 0;

QY 2 SNAATSYINAIINISKVSVD 20
Db 114 NGAVTSYINTIVNSKSITE 132

RESULT 15
O97N55 STRPN
ID Q97N55_STRPN PRELIMINARY; PRT; 392 AA.
AC Q97N55;
DT 01-OCT-2001 (TrEMBLrel. 18, Created)
DT 01-OCT-2001 (TrEMBLrel. 18, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Secreted 45 kd protein.
GN OrderedLocustNames-SP2216;
OS Streptococcus pneumoniae.
OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
OC Streptococcus.
OX NCBI_TaxID=1313;
RN [1]_
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=ATCC BAA-334 / TIGR4;
RX MEDLINE=21357209; PubMed=11463916; DOI=10.1126/science.1061217;
RA Tetterlin H., Nelson K.E., Paulsen I.T., Eisen J.A., Read T.D.,
RA Peterson S.N., Heidelberg J.F., DeBoy R.T., Haft D.H., Dodson R.J.,
RA Durkin A.S., Gwinn M.L., Kolonay J.F., Nelson W.C., Peterson J.D.,
RA Unayam L.A., White O., Salzberg S.L., Lewis M.R., Radune D.,
RA Holtzapple E.K., Knouri H.M., Wolf A.M., Utterback T.R., Hansen C.L.,
RA McDonald L.A., Feldblyum T.V., Angiuoli S.V., Dickinson T.,
RA Hickey E.K., Holt I.E., Loftus B.J., Yang F., Smith H.O., Venter J.C.,
RA Dougherty B.A., Morrison D.A., Hollingshead S.K., Fraser C.M.;
RT "Complete genome sequence of a virulent isolate of Streptococcus
RT pneumoniae."
RL Science 293:498-506(2001).
DR EMBL; AE007509; AAK76264.1; -; Genomic_DNA.
DR PIR; B98124; B98124.
DR PIR; G95258; G95258.
DR TIGR; SP2216; -.
DR InterPro; IPR007921; CHAP.
DR InterPro; IPR009148; SIDA.
DR Pfam; PF05257; CHAP; 1.
DR PRINTS; PR01852; SIBAPROTEIN.
DR PROSITE; PS0911; CHAP; 1.
KW Complete proteome.
SQ SEQUENCE 392 AA; 41697 MW; B5A5A860FAEA77DD CRC64;

Query Match 67.0%; Score 63; DB 2; Length 392;
Best Local Similarity 57.9%; Pred. No. 0.077;
Matches 11; Conservative 5; Mismatches 3; Indels 0; Gaps 0;

QY 2 SNAATSYINAIINISKVSVD 20
Db 114 NGAVTSYINTIVNSKSITE 132

```

GenCore version 5.1.6
Copyright (c) 1993 - 2006 Compugen Ltd.

OM protein - protein search, using sw model

Run on: January 25, 2006, 18:50:05 ; Search time 16.6 Seconds
(without alignments)
99.609 Million cell updates/sec

Title: US-10-797-821-1
Perfect score: 94
Sequence: 1 KSNATSYINAINSKSVSD 20

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 572060 seqs, 82675679 residues

Total number of hits satisfying chosen parameters: 572060

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued Patents AA:*
1: /cgn2_6/prodata/1/iaa/5.COMB.pep:*
2: /cgn2_6/prodata/1/iaa/6.COMB.pep:*
3: /cgn2_6/prodata/1/iaa/H.COMB.pep:*
4: /cgn2_6/prodata/1/iaa/PCTUS.COMB.pep:*
5: /cgn2_6/prodata/1/iaa/RE.COMB.pep:*
6: /cgn2_6/prodata/1/iaa/backfiles1.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	63	67.0	210	2 US-09-222-938A-67	Sequence 67, Appl
2	63	67.0	392	2 US-09-583-110-4374	Sequence 4374, Ap
3	63	67.0	399	2 US-09-107-433-3230	Sequence 3230, Ap
4	59	62.8	461	1 US-08-186-222-2	Sequence 2, Appli
5	52	55.3	525	2 US-09-107-532A-5095	Sequence 5095, Ap
6	45	47.9	152	1 US-08-243-082-24	Sequence 24, Appl
7	45	47.9	172	1 US-08-243-082-18	Sequence 18, Appl
8	45	47.9	808	2 US-09-248-796A-18679	Sequence 18679, A
9	44	46.8	331	2 US-09-094-557-34	Sequence 34, Appl
10	44	46.8	646	2 US-09-625-188-10	Sequence 30, Appl
11	44	46.8	910	2 US-09-902-540-10793	Sequence 10793, A
12	44	46.8	946	2 US-09-657-931A-10	Sequence 10, Appl
13	43	45.7	299	2 US-09-543-681A-5625	Sequence 5625, Ap
14	43	45.7	815	2 US-09-248-796A-19069	Sequence 19069, A
15	42	44.7	170	2 US-09-708-392-15	Sequence 15, Appl
16	42	44.7	170	2 US-09-949-016-6352	Sequence 6352, Ap
17	42	44.7	170	2 US-09-949-016-6352	Sequence 17, Appl
18	42	44.7	192	2 US-09-949-016-7555	Sequence 7555, Ap
19	42	44.7	211	2 US-09-248-796A-18107	Sequence 18107, A
20	42	44.7	591	2 US-09-134-000C-5327	Sequence 5327, Ap
21	41.5	44.1	203	2 US-09-538-092-11	Sequence 11, Appl
22	41	43.6	508	1 US-08-724-281-1	Sequence 1, Appli
23	41	43.6	508	2 US-09-134-377-1	Sequence 1, Appli
24	41	43.6	508	2 US-10-665-537-1	Sequence 1, Appli
25	41	43.6	948	2 US-09-248-796A-19297	Sequence 19297, A
26	40	42.6	146	2 US-09-373-750-3	Sequence 3, Appli
27	40	42.6	249	2 US-09-413-814-100	Sequence 100, App

28	40	42.6	311	2 US-09-489-039A-13109	Sequence 13109, A
29	40	42.6	500	1 US-08-704-398-2	Sequence 2, Appli
30	40	42.6	500	4 PCT-US95-05966-2	Sequence 2, Appli
31	40	42.6	537	2 US-09-248-796A-19378	Sequence 19378, A
32	40	42.6	756	2 US-09-248-796A-14799	Sequence 14799, A
33	40	42.6	919	2 US-09-270-767-44725	Sequence 44725, A
34	39.5	42.0	195	2 US-09-222-939-29	Sequence 29, Appl
35	39.5	42.0	195	2 US-10-023-528-29	Sequence 29, Appl
36	39.5	42.0	195	2 US-10-423-330-29	Sequence 29, Appl
37	39	41.5	30	1 US-08-243-082-3	Sequence 3, Appli
38	39	41.5	31	1 US-08-243-082-5	Sequence 5, Appli
39	39	41.5	206	2 US-09-830-230A-294	Sequence 294, App
40	39	41.5	221	2 US-09-308-003-50	Sequence 50, Appl
41	39	41.5	241	2 US-09-353-242A-1	Sequence 1, Appli
42	39	41.5	250	2 US-09-830-230A-293	Sequence 293, App
43	39	41.5	366	2 US-09-787-069-2	Sequence 2, Appli
44	39	41.5	388	2 US-09-308-003-16	Sequence 16, Appl
45	39	41.5	442	2 US-09-540-236-3378	Sequence 3378, Ap

ALIGNMENTS

RESULT 1
US-09-222-938A-67
; Sequence 67, Application US/09222938A
; Patent No. 6437108
; GENERAL INFORMATION:
; APPLICANT: Youngman, Philip
; APPLICANT: Pritz, Christian
; APPLICANT: Murphy, Christopher
; APPLICANT: Guzman, Luz-Maria
; TITLE OF INVENTION: ESSENTIAL BACTERIAL GENES AND THEIR USE
; FILE REFERENCE: 07334/060001
; CURRENT APPLICATION NUMBER: US/09/222,938A
; CURRENT FILING DATE: 1998-12-30
; NUMBER OF SEQ ID NOS: 102
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 67
; LENGTH: 210
; TYPE: PRT
; ORGANISM: Streptococcus pneumoniae
US-09-222-938A-67

Query Match 67.0%; Score 63; DB 2; Length 210;
Best Local Similarity 57.9%; Pred. No. 0.026;
Matches 11; Conservative 5; Mismatches 3; Indels 0; Gaps 0;

QY 2 SNAATSYINAINSKSVSD 20
: ||||| :|||:
Db 114 NGAVTSYINTIVNSKSITE 132

RESULT 2

US-09-583-110-4374
; Sequence 4374, Application US/09583110
; Patent No. 6699703
; GENERAL INFORMATION:
; APPLICANT: Lynn Doucette-Staum et al.
; TITLE OF INVENTION: Nucleic Acid and Amino Acid Sequences Relating to Streptococcus
; FILE REFERENCE: PATH00-07A
; CURRENT APPLICATION NUMBER: US/09/583,110
; CURRENT FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: US 09/107,433
; PRIOR FILING DATE: 1998-06-30
; PRIOR APPLICATION NUMBER: US 60/085,131
; PRIOR FILING DATE: 1998-05-12
; PRIOR APPLICATION NUMBER: US 60/051,553
; PRIOR FILING DATE: 1997-07-02
; NUMBER OF SEQ ID NOS: 5322
; SEQ ID NO 4374
; LENGTH: 392

```
; TYPE: PRT
; ORGANISM: Streptococcus pneumoniae
US-09-583-110-4374

Query Match      67.0%; Score 63; DB 2; Length 392;
Best Local Similarity 57.9%; Pred. No. 0.053;
Matches 11; Conservative 5; Mismatches 3; Indels 0; Gaps 0;

QY      2 SNAATSYINAINSKSVSD 20
Db      114 NGAVTSYINTIVNSKSITE 132

RESULT 3
US-09-107-433-3230
; Sequence 3230, Application US/09107433
; Patent No. 6800744
; GENERAL INFORMATION:
; APPLICANT: Lynn A Doucette-Stamm and David Bush
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID
; SEQUENCES RELATING TO STREPTOCOCCUS PNEUMONIAE FOR DIAGNOSTICS AND THERAPEUTICS
; NUMBER OF SEQUENCES: 5206
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: GENOME THERAPEUTICS CORPORATION
; STREET: 100 Beaver Street
; CITY: Waltham
; STATE: Massachusetts
; COUNTRY: USA
; ZIP: 02354
; COMPUTER READABLE FORM:
; MEDIUM TYPE: CD/ROM ISO9660
; COMPUTER: <Unknown>
; OPERATING SYSTEM: <Unknown>
; SOFTWARE: <Unknown>
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/107,433
; FILING DATE: 30-Jun-1998
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 60/ 085131
; FILING DATE: May 12, 1998
; APPLICATION NUMBER: 60/051553
; FILING DATE: July 2, 1997
; ATTORNEY/AGENT INFORMATION:
; NAME: Ariniello, Pamela Deneke
; REGISTRATION NUMBER: 40,489
; REFERENCE/DOCKET NUMBER: GTC-011
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (781)893-5007
; TELEFAX: (781)893-8277
; INFORMATION FOR SEQ ID NO: 3230:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 399 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; HYPOTHETICAL: YES
; ORIGINAL SOURCE:
; ORGANISM: Streptococcus pneumoniae
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (B) LOCATION 1...399
; SEQUENCE DESCRIPTION: SEQ ID NO: 3230:
US-09-107-433-3230

Query Match      67.0%; Score 63; DB 2; Length 399;
Best Local Similarity 57.9%; Pred. No. 0.054;
Matches 11; Conservative 5; Mismatches 3; Indels 0; Gaps 0;

QY      2 SNAATSYINAINSKSVSD 20
Db      121 NGAVTSYINTIVNSKSITE 139

; TYPE: PRT
; ORGANISM: Streptococcus pneumoniae
US-08-186-222-2
; Sequence 2, Application US/08186222
; Patent No. 5559007
; GENERAL INFORMATION:
; APPLICANT: Suri, Bruno
; APPLICANT: Schmitz, Albert
; TITLE OF INVENTION: Bacterial Vectors
; NUMBER OF SEQUENCES: 4
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: CIBA-GEIGY Corporation
; STREET: 7 Skyline Drive
; CITY: Hawthorne
; STATE: New York
; COUNTRY: USA
; ZIP: 10532
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/186,222
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/672,205
; FILING DATE: 19-MAR-1991
; APPLICATION NUMBER: GB 9006400.7
; FILING DATE: 22-MAR-1990
; ATTORNEY/AGENT INFORMATION:
; NAME: Villamizar, Joann
; REGISTRATION NUMBER: 30,598
; REFERENCE/DOCKET NUMBER: 4-17994/A
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (914)785-7121
; TELEFAX: (914)347-5769
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 461 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-186-222-2

Query Match      62.8%; Score 59; DB 1; Length 461;
Best Local Similarity 47.4%; Pred. No. 0.27;
Matches 9; Conservative 10; Mismatches 0; Indels 0; Gaps 0;

QY      2 SNAATSYINAINSKSVSD 20
Db      113 NSSATNYMDAVVNSKSLTD 131

RESULT 5
US-09-107-532A-5095
; Sequence 5095, Application US/09107532A
; Patent No. 6583275
; GENERAL INFORMATION:
; APPLICANT: Lynn A Doucette-Stamm and David Bush
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: GENOME THERAPEUTICS CORPORATION
; STREET: 100 Beaver Street
; CITY: Waltham
; STATE: Massachusetts
; COUNTRY: USA
; ZIP: 02354
; COMPUTER READABLE FORM:
; MEDIUM TYPE: CD/ROM ISO9660
```



```

; COMPUTER: PC
; OPERATING SYSTEM: <Unknown>
; SOFTWARE: ASCII
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/107,532A
; FILING DATE: 30-JUN-1998
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 60/085,598
; FILING DATE: 14 May 1998
; APPLICATION NUMBER: 60/051571
; FILING DATE: July 2, 1997
; ATTORNEY/AGENT INFORMATION:
; NAME: Ariniello, Pamela Deneke
; REGISTRATION NUMBER: 40,489
; REFERENCE/DOCKET NUMBER: GTC-012
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (781)893-5007
; TELEFAX: (781)893-8277
; INFORMATION FOR SEQ ID NO: 5095:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 525 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; HYPOTHETICAL: YES
; ORIGINAL SOURCE:
; ORGANISM: Enterococcus faecium
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (B) LOCATION 1...525
; SEQUENCE DESCRIPTION: SEQ ID NO: 5095:
US-09-107-532A-5095

Query Match 55.3%; Score 52; DB 2; Length 525;
Best Local Similarity 42.1%; Pred. No. 3.9;
Matches 8; Conservative 9; Mismatches 2; Indels 0; Gaps 0;

QY 2 SNAATSYINAINSKSVSD 20
||::||::||::||
DB 115 SNTSSNYIDAVLNADSLAD 133

RESULT 6
US-08-243-082-24
; Sequence 24, Application US/08243082
; Patent No. 5506120
; GENERAL INFORMATION:
; APPLICANT: YAMAMOTO, Hiroaki
; APPLICANT: YAMASHITA, Kunihiko
; TITLE OF INVENTION: METHOD OF PRODUCING PEPTIDES OR
; NUMBER OF SEQUENCES: 26
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Spencer, Frank & Schneider
; STREET: 1111 Nineteenth Street, N.W.
; CITY: Washington
; STATE: D.C.
; COUNTRY: U.S.A.
; ZIP: 20036
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/243,082
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/07/853,754
; FILING DATE: 05-JUN-1992
; ATTORNEY/AGENT INFORMATION:
; NAME: Schneller, John W.
; REGISTRATION NUMBER: 26,031
; REFERENCE/DOCKET NUMBER: KUWAT 0010
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202) 828-8000
; TELEFAX: (202) 828-8038
; TELETYPE: SPENCER 64267
; INFORMATION FOR SEQ ID NO: 18:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 172 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; FEATURE:
; NAME/KEY: Modified-site
; LOCATION: <1
; OTHER INFORMATION: /note= "Protein A(1-402)"
US-08-243-082-18

```

```

; REGISTRATION NUMBER: 26,031
; REFERENCE/DOCKET NUMBER: KUWAT 0010
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202) 828-8000
; TELEFAX: (202) 828-8038
; TELETYPE: SPENCER 64267
; INFORMATION FOR SEQ ID NO: 24:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 152 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
US-08-243-082-24

Query Match 47.9%; Score 45; DB 1; Length 152;
Best Local Similarity 45.0%; Pred. No. 12;
Matches 9; Conservative 3; Mismatches 8; Indels 0; Gaps 0;

QY 1 KSNATSYINAINSKSVSD 20
||::||::||::||
DB 45 KQWAVKKYLSINLGRHSD 64

RESULT 7
US-08-243-082-18
; Sequence 18, Application US/08243082
; Patent No. 5506120
; GENERAL INFORMATION:
; APPLICANT: YAMAMOTO, Hiroaki
; APPLICANT: YAMASHITA, Kunihiko
; TITLE OF INVENTION: METHOD OF PRODUCING PEPTIDES OR
; NUMBER OF SEQUENCES: 26
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Spencer, Frank & Schneider
; STREET: 1111 Nineteenth Street, N.W.
; CITY: Washington
; STATE: D.C.
; COUNTRY: U.S.A.
; ZIP: 20036
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/243,082
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/07/853,754
; FILING DATE: 05-JUN-1992
; ATTORNEY/AGENT INFORMATION:
; NAME: Schneller, John W.
; REGISTRATION NUMBER: 26,031
; REFERENCE/DOCKET NUMBER: KUWAT 0010
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202) 828-8000
; TELEFAX: (202) 828-8038
; TELETYPE: SPENCER 64267
; INFORMATION FOR SEQ ID NO: 18:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 172 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; FEATURE:
; NAME/KEY: Modified-site
; LOCATION: <1
; OTHER INFORMATION: /note= "Protein A(1-402)"
US-08-243-082-18

```

```

Query Match      47.9%; Score 45; DB 1; Length 172;
Best Local Similarity 45.0%; Pred. No. 14;
Matches 9; Conservative 3; Mismatches 8; Indels 0; Gaps 0;

QY      1 KSNAAATSYINAINSKVSVD 20
Db      34 KOMAVKKYLSINLNGKRHSD 53

RESULT 8
US-09-248-796A-18679
; Sequence 18679, Application US/09248796A
; Patent No. 6747137
; GENERAL INFORMATION:
; APPLICANT: Keith Weinstock et al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO CANDIDA ALBICAN
; FILE REFERENCE: 107196.132
; CURRENT APPLICATION NUMBER: US/09/248,796A
; PRIOR FILING DATE: 1999-02-12
; PRIOR APPLICATION NUMBER: US 60/074,725
; PRIOR FILING DATE: 1998-02-13
; PRIOR APPLICATION NUMBER: US 60/096,409
; PRIOR FILING DATE: 1998-08-13
; NUMBER OF SEQ ID NOS: 28208
; SEQ ID NO 18679
; LENGTH: 808
; TYPE: PRT
; ORGANISM: Candida albicans
US-09-248-796A-18679

Query Match      47.9%; Score 45; DB 2; Length 808;
Best Local Similarity 44.4%; Pred. No. 80;
Matches 8; Conservative 4; Mismatches 6; Indels 0; Gaps 0;

QY      3 NAAATSYINAINSKVSVD 20
Db      637 NAAQTVLNLVINDKDILD 654

RESULT 9
US-09-094-557-34
; Sequence 34, Application US/09094557
; Patent No. 6225531
; GENERAL INFORMATION:
; APPLICANT: Kakitani, Makoto
; APPLICANT: Umamoto, Naoyuki
; APPLICANT: Ishida, Isao
; APPLICANT: Iwamatsu, Akihiro
; APPLICANT: Yoshikawa, Masaaki
; APPLICANT: Yamaoka, Naoto
; TITLE OF INVENTION: GLUCAN ELICITOR RECEPTOR, DNA MOLECULE
; TITLE OF INVENTION: CODING THEREFOR, FUNGUS-RESISTANT PLANTS TRANSFORMED WITH
; TITLE OF INVENTION: THE DNA MOLECULE AND METHOD FOR CREATING THE PLANTS
; NUMBER OF SEQUENCES: 36
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Foley & Lardner
; STREET: 3000 K Street, N.W., Suite 500
; CITY: Washington
; STATE: D.C.
; COUNTRY: USA
; ZIP: 20007-5109
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/094,557
; FILING DATE: 15-JUN-1998
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:

```

```

; APPLICATION NUMBER: PCT/JP96/03653
; FILING DATE: 13-DEC-1996
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: JP 136100/1994
; FILING DATE: 17-JUN-1994
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: JP 347823/1995
; FILING DATE: 15-DEC-1995
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/591,566
; FILING DATE: 14-FEB-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: Bent, Stephen A.
; REGISTRATION NUMBER: 29,768
; REFERENCE/DOCKET NUMBER: 081356/0116
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202)672-5300
; TELEFAX: (202)672-5399
; TELEX: 904136
; INFORMATION FOR SEQ ID NO: 34:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 331 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-09-094-557-34

Query Match      46.8%; Score 44; DB 2; Length 331;
Best Local Similarity 69.2%; Pred. No. 42;
Matches 9; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY      2 SNAATSYINAIN 14
Db      170 SNSASSYITPIIN 182

RESULT 10
US-09-625-188-10
; Sequence 10, Application US/09625188
; Patent No. 6307037
; GENERAL INFORMATION:
; APPLICANT: No. 6307037artis AG
; TITLE OF INVENTION: Fungal Target Genes and Methods
; FILE REFERENCE: PB/5-31285P1
; CURRENT APPLICATION NUMBER: US/09/625,188
; CURRENT FILING DATE: 2000-07-21
; NUMBER OF SEQ ID NOS: 44
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 10
; LENGTH: 646
; TYPE: PRT
; ORGANISM: Ashbya gossypii
US-09-625-188-10

Query Match      46.8%; Score 44; DB 2; Length 646;
Best Local Similarity 42.1%; Pred. No. 89;
Matches 8; Conservative 5; Mismatches 6; Indels 0; Gaps 0;

QY      1 KSNAAATSYINAINSKVS 19
Db      625 RSNNTDLLNQLNGKKVT 643

RESULT 11
US-09-902-540-10793
; Sequence 10793, Application US/09902540
; Patent No. 6833447
; GENERAL INFORMATION:
; APPLICANT: Goldman, Barry S.
; APPLICANT: Hinkle, Gregory J.
; APPLICANT: Slater, Steven C.
; APPLICANT: Wiegand, Roger C.
; TITLE OF INVENTION: Myxococcus xanthus Genome Sequences and Uses Thereof

```

; FILE REFERENCE: 38-10(15849)B
; CURRENT APPLICATION NUMBER: US/09/902,540
; CURRENT FILING DATE: 2001-07-10
; PRIOR APPLICATION NUMBER: 60/217,883
; PRIOR FILING DATE: 2000-07-10
; NUMBER OF SEQ ID NOS: 16825
; SEQ ID NO 10793
; LENGTH: 910
; TYPE: PRT
; ORGANISM: Myxococcus xanthus
US-09-902-540-10793

Query Match 46.8%; Score 44; DB 2; Length 910;
Best Local Similarity 71.4%; Pred. No. 1.3e+02;
Matches 10; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 2 SNAATSYINAININS 15
Db 661 SNAITTTINAIIDS 674
||| |::: ||| |::: |

RESULT 12

US-09-657-931A-10
; Sequence 10, Application US/09657931A
; Patent No. 6586197
; GENERAL INFORMATION:
; APPLICANT: ADANG, MICHAEL J
; APPLICANT: LUO, KE
; TITLE OF INVENTION: METHODS AND MATERIALS FOR IDENTIFYING NOVEL PESTICIDE AGENTS
; FILE REFERENCE: UGR-101X
; CURRENT APPLICATION NUMBER: US/09/657,931A
; CURRENT FILING DATE: 2000-09-07
; NUMBER OF SEQ ID NOS: 15
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 10
; LENGTH: 946
; TYPE: PRT
; ORGANISM: Plutella xylostella
US-09-657-931A-10

Query Match 46.8%; Score 44; DB 2; Length 946;
Best Local Similarity 41.2%; Pred. No. 1.4e+02;
Matches 7; Conservative 7; Mismatches 3; Indels 0; Gaps 0;

QY 2 SNAATSYINAININSKV 18
Db 656 ANAAFYLNVLSSGV 672
: ||| |::: ||| |::: |

RESULT 13

US-09-543-681A-5625
; Sequence 5625, Application US/09543681A
; Patent No. 6605709
; GENERAL INFORMATION:
; APPLICANT: GARY BRETON
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PROTEUS MIRABILIS
; FILE REFERENCE: 2709.1002-001
; CURRENT APPLICATION NUMBER: US/09/543,681A
; CURRENT FILING DATE: 2000-04-05
; PRIOR APPLICATION NUMBER: US 60/128,706
; PRIOR FILING DATE: 1999-04-09
; NUMBER OF SEQ ID NOS: 8344
; SEQ ID NO 5625
; LENGTH: 299
; TYPE: PRT
; ORGANISM: Proteus mirabilis
US-09-543-681A-5625

Query Match 45.7%; Score 43; DB 2; Length 299;
Best Local Similarity 42.1%; Pred. No. 55;
Matches 8; Conservative 7; Mismatches 4; Indels 0; Gaps 0;

QY 2 SNAATSYINAININSKVSVD 20
Db 21 SDSSTSFLESIINSKRGD 39
|::: |::: ||| |::: |

RESULT 14

US-09-248-796A-19069
; Sequence 19069, Application US/09248796A
; Patent No. 6747137
; GENERAL INFORMATION:
; APPLICANT: Keith Weinstock et al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO CANDIDA ALBICA.
; FILE REFERENCE: 107196.132
; CURRENT APPLICATION NUMBER: US/09/248,796A
; CURRENT FILING DATE: 1999-02-12
; PRIOR APPLICATION NUMBER: US 60/074,725
; PRIOR FILING DATE: 1998-02-13
; PRIOR APPLICATION NUMBER: US 60/096,409
; PRIOR FILING DATE: 1998-08-13
; NUMBER OF SEQ ID NOS: 28208
; SEQ ID NO 19069
; LENGTH: 815
; TYPE: PRT
; ORGANISM: Candida albicans
US-09-248-796A-19069

Query Match 45.7%; Score 43; DB 2; Length 815;
Best Local Similarity 53.3%; Pred. No. 1.7e+02;
Matches 8; Conservative 3; Mismatches 4; Indels 0; Gaps 0;

QY 4 AATSYINAININSKV 18
Db 653 AGTSYLYALYNSKEI 667
| ||| |::: ||| |::: |

RESULT 15

US-09-708-392-15
; Sequence 15, Application US/09708392
; Patent No. 6734186
; GENERAL INFORMATION:
; APPLICANT: Pfizer Inc
; APPLICANT: Pfizer Limited
; APPLICANT: Wayman, Chris
; APPLICANT: Maw, G
; TITLE OF INVENTION: Pharmaceutical
; FILE REFERENCE: PC10343AAKM
; CURRENT APPLICATION NUMBER: US/09/708,392
; CURRENT FILING DATE: 2000-11-08
; PRIOR APPLICATION NUMBER: GB 9926437.6
; PRIOR FILING DATE: 1999-11-08
; PRIOR APPLICATION NUMBER: GB 0004021.2
; PRIOR FILING DATE: 2000-02-18
; PRIOR APPLICATION NUMBER: GB 0013001.3
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: GB 0016563.9
; PRIOR FILING DATE: 2000-07-05
; PRIOR APPLICATION NUMBER: GB 0017141.3
; PRIOR FILING DATE: 2000-07-12
; PRIOR APPLICATION NUMBER: US 60/175,161
; PRIOR FILING DATE: 2000-01-07
; PRIOR APPLICATION NUMBER: US 60/192,962
; PRIOR FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: US 60/217,479
; PRIOR FILING DATE: 2000-07-11
; PRIOR APPLICATION NUMBER: US 60/221,014
; PRIOR FILING DATE: 2000-07-27
; PRIOR APPLICATION NUMBER: US 60/221,093
; PRIOR FILING DATE: 2000-07-27
; NUMBER OF SEQ ID NOS: 20
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 15
; LENGTH: 170

```

; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-708-392-15

```

Query Match 44.7%; Score 42; DB 2; Length 170;
Best Local Similarity 40.0%; Pred. No. 42;
Matches 8; Conservative 4; Mismatches 8; Indels 0; Gaps 0;

Qy	1	K	S	N	A	T	S	Y	I	N	A	I	N	S	K	S	V	S	D	20
																			:	
Db	139	K	O	M	A	V	K	K	Y	L	N	S	I	L	N	G	K	R	S	158

Search completed: January 25, 2006, 19:18:11
Job time : 16.6 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2006 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: January 25, 2006, 19:13:51 ; Search time 60 Seconds
(without alignments)
139.276 Million cell updates/sec

Title: US-10-797-821-1

Perfect score: 94

Sequence: 1 KSNAAATSYINAIINAKSVSD 20

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1867569 seqs, 417829326 residues

Total number of hits satisfying chosen parameters: 1867569

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications AA Main:

- 1: /cgn2_6/prodata/1/pubpaa/US07_PUBCOMB.pap.*
- 2: /cgn2_6/prodata/1/pubpaa/US08_PUBCOMB.pap.*
- 3: /cgn2_6/prodata/1/pubpaa/US09_PUBCOMB.pap.*
- 4: /cgn2_6/prodata/1/pubpaa/US10A_PUBCOMB.pap.*
- 5: /cgn2_6/prodata/1/pubpaa/US10B_PUBCOMB.pap.*
- 6: /cgn2_6/prodata/1/pubpaa/US11_PUBCOMB.pap.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	94	100.0	20	4	US-10-383-930-1
2	94	100.0	20	5	US-10-797-821-1
3	94	100.0	431	4	US-10-383-930-29
4	94	100.0	431	4	US-10-383-930-30
5	94	100.0	431	4	US-10-383-930-33
6	94	100.0	431	5	US-10-797-821-29
7	94	100.0	431	5	US-10-797-821-30
8	94	100.0	431	5	US-10-797-821-33
9	94	100.0	432	4	US-10-383-930-31
10	94	100.0	432	4	US-10-383-930-32
11	94	100.0	432	5	US-10-797-821-31
12	94	100.0	432	5	US-10-797-821-32
13	75	79.8	20	4	US-10-383-930-11
14	75	79.8	20	5	US-10-797-821-11
15	69	73.4	398	5	US-10-474-792-600
16	63	67.0	210	4	US-10-154-251-67
17	63	67.0	392	5	US-10-472-928-4652
18	63	67.0	399	5	US-10-617-320-3230
19	52	55.3	524	4	US-10-282-122A-57658
20	47	50.0	150	5	US-10-480-456-8
21	47	50.0	2139	5	US-10-480-456-1
22	44	46.8	91	4	US-10-282-122A-45418
23	44	46.8	331	3	US-09-816-127-34
24	44	46.8	504	4	US-10-282-122A-71895
25	43	45.7	126	4	US-10-425-115-225682
26	43	45.7	148	4	US-10-424-599-228655
27	43	45.7	860	5	US-10-741-849-7218

28 42.5 45.2 287 4 US-10-437-963-196528 Sequence 196528,
29 42 44.7 30 5 US-10-851-691-122 Sequence 122, App
30 42 44.7 30 5 US-10-892-981A-126 Sequence 126, App
31 42 44.7 64 4 US-10-425-115-218215 Sequence 218215,
32 42 44.7 170 5 US-10-686-390-15 Sequence 15, Appl
33 42 44.7 170 5 US-10-686-390-17 Sequence 17, Appl
34 42 44.7 170 5 US-10-686-282-15 Sequence 15, Appl
35 42 44.7 170 5 US-10-686-282-17 Sequence 17, Appl
36 42 44.7 170 5 US-10-686-349-15 Sequence 17, Appl
37 42 44.7 170 5 US-10-686-349-17 Sequence 17, Appl
38 42 44.7 218 4 US-10-032-585-7853 Sequence 7853, Ap
39 42 44.7 371 4 US-10-767-701-43218 Sequence 43218, A
40 42 44.7 552 4 US-10-369-493-10160 Sequence 10160, A
41 42 44.7 559 4 US-10-282-122A-47206 Sequence 47206, A
42 42 44.7 748 4 US-10-369-493-8217 Sequence 8217, Ap
43 41 43.6 30 5 US-10-851-691-130 Sequence 130, App
44 41 43.6 30 5 US-10-892-981A-134 Sequence 134, App
45 41 43.6 50 4 US-10-767-701-57338 Sequence 57338, A

ALIGNMENTS

RESULT 1

US-10-383-930-1

; Sequence 1, Application US/10383930

; Publication NO. US20040127400A1

; GENERAL INFORMATION:

; APPLICANT: Smith, Daniel J

; APPLICANT: Taubman, Martin A

; TITLE OF INVENTION: Immunogenicity of Glucan Binding Protein

; FILE REFERENCE: 25669-018

; CURRENT APPLICATION NUMBER: US/10/383,930

; CURRENT FILING DATE: 2003-03-07

; PRIOR APPLICATION NUMBER: 60/402,483

; PRIOR FILING DATE: 2002-08-08

; PRIOR APPLICATION NUMBER: 60/363,209

; PRIOR FILING DATE: 2002-03-07

; NUMBER OF SEQ ID NOS: 41

; SOFTWARE: PatentIn version 3.2

; SEQ ID NO 1

; LENGTH: 20

; TYPE: PRT

; ORGANISM: Streptococcus mutans

US-10-383-930-1

Query Match 100.0%; Score 94; DB 4; Length 20;
Best Local Similarity 100.0%; Pred. No. 4.2e-08; Mismatches 0; Indels 0; Gaps 0;
Matches 20; Conservative 0;

QY 1 KSNAAATSYINAIINAKSVSD 20

|||||

Db 1 KSNAAATSYINAIINAKSVSD 20

RESULT 2

US-10-797-821-1

; Sequence 1, Application US/10797821

; Publication No. US20050031633A1

; GENERAL INFORMATION:

; APPLICANT: Smith, Daniel J.

; APPLICANT: Taubman, Martin A.

; TITLE OF INVENTION: Glucan Binding Protein and Glycosyltransferase Immunogens

; FILE REFERENCE: 25669-020

; CURRENT APPLICATION NUMBER: US/10/797,821

; CURRENT FILING DATE: 2004-03-09

; PRIOR APPLICATION NUMBER: 10/383,930

; PRIOR FILING DATE: 2003-03-07

; PRIOR APPLICATION NUMBER: 60/363,209

; PRIOR FILING DATE: 2002-03-07

; PRIOR APPLICATION NUMBER: 60/402,483

; PRIOR FILING DATE: 2002-08-08

; PRIOR APPLICATION NUMBER: 09/290,049

; PRIOR FILING DATE: 1999-04-12
; PRIOR APPLICATION NUMBER: 60/081,550
; PRIOR FILING DATE: 1998-04-13
; PRIOR APPLICATION NUMBER: 60/115,142
; PRIOR FILING DATE: 1999-01-08
; NUMBER OF SEQ ID NOS: 45
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 1
; LENGTH: 20
; TYPE: PRT
; ORGANISM: Artificial
; FEATURE:
; OTHER INFORMATION: Gpbb-derived MHC class II (SYI) peptide
US-10-797-821-1

Query Match 100.0%; Score 94; DB 5; Length 20;
Best Local Similarity 100.0%; Pred. No. 4.2e-08;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KSNAAATSYINAIINSKVSVD 20
|||||
Db 1 KSNAAATSYINAIINSKVSVD 20

RESULT 3

US-10-383-930-29
; Sequence 29, Application US/10383930
; Publication No. US20040127400A1
; GENERAL INFORMATION:
; APPLICANT: Smith, Daniel J
; APPLICANT: Taubman, Martin A
; TITLE OF INVENTION: Immunogenicity of Glucan Binding Protein
; FILE REFERENCE: 25669-018
; CURRENT APPLICATION NUMBER: US/10/383,930
; CURRENT FILING DATE: 2003-03-07
; PRIOR APPLICATION NUMBER: 60/402,483
; PRIOR FILING DATE: 2002-08-08
; PRIOR APPLICATION NUMBER: 60/363,209
; PRIOR FILING DATE: 2002-03-07
; NUMBER OF SEQ ID NOS: 41
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 29
; LENGTH: 431
; TYPE: PRT
; ORGANISM: Streptococcus mutans
US-10-383-930-29

Query Match 100.0%; Score 94; DB 4; Length 431;
Best Local Similarity 100.0%; Pred. No. 1.3e-06;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KSNAAATSYINAIINSKVSVD 20
|||||
Db 113 KSNAAATSYINAIINSKVSVD 132

RESULT 4

US-10-383-930-30
; Sequence 30, Application US/10383930
; Publication No. US20040127400A1
; GENERAL INFORMATION:
; APPLICANT: Smith, Daniel J
; APPLICANT: Taubman, Martin A
; TITLE OF INVENTION: Immunogenicity of Glucan Binding Protein
; FILE REFERENCE: 25669-018
; CURRENT APPLICATION NUMBER: US/10/383,930
; CURRENT FILING DATE: 2003-03-07
; PRIOR APPLICATION NUMBER: 60/402,483
; PRIOR FILING DATE: 2002-08-08
; PRIOR APPLICATION NUMBER: 60/363,209
; PRIOR FILING DATE: 2002-03-07
; NUMBER OF SEQ ID NOS: 41
; SOFTWARE: PatentIn version 3.2

; SEQ ID NO 30
; LENGTH: 431
; TYPE: PRT
; ORGANISM: Streptococcus mutans
US-10-383-930-30

Query Match 100.0%; Score 94; DB 4; Length 431;
Best Local Similarity 100.0%; Pred. No. 1.3e-06;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KSNAAATSYINAIINSKVSVD 20
|||||
Db 113 KSNAAATSYINAIINSKVSVD 132

RESULT 5

US-10-383-930-33
; Sequence 33, Application US/10383930
; Publication No. US20040127400A1
; GENERAL INFORMATION:
; APPLICANT: Smith, Daniel J
; APPLICANT: Taubman, Martin A
; TITLE OF INVENTION: Immunogenicity of Glucan Binding Protein
; FILE REFERENCE: 25669-018
; CURRENT APPLICATION NUMBER: US/10/383,930
; CURRENT FILING DATE: 2003-03-07
; PRIOR APPLICATION NUMBER: 60/402,483
; PRIOR FILING DATE: 2002-08-08
; PRIOR APPLICATION NUMBER: 60/363,209
; PRIOR FILING DATE: 2002-03-07
; NUMBER OF SEQ ID NOS: 41
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 33
; LENGTH: 431
; TYPE: PRT
; ORGANISM: Streptococcus mutans
US-10-383-930-33

Query Match 100.0%; Score 94; DB 4; Length 431;
Best Local Similarity 100.0%; Pred. No. 1.3e-06;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KSNAAATSYINAIINSKVSVD 20
|||||
Db 113 KSNAAATSYINAIINSKVSVD 132

RESULT 6

US-10-797-821-29
; Sequence 29, Application US/10797821
; Publication No. US20050031633A1
; GENERAL INFORMATION:
; APPLICANT: Smith, Daniel J.
; APPLICANT: Taubman, Martin A.
; TITLE OF INVENTION: Glucan Binding Protein and Glycosyltransferase Immunogens
; FILE REFERENCE: 25669-020
; CURRENT APPLICATION NUMBER: US/10/797,821
; CURRENT FILING DATE: 2004-03-09
; PRIOR APPLICATION NUMBER: 10/383,930
; PRIOR FILING DATE: 2003-03-07
; PRIOR APPLICATION NUMBER: 60/363,209
; PRIOR FILING DATE: 2002-03-07
; PRIOR APPLICATION NUMBER: 60/402,483
; PRIOR FILING DATE: 2002-08-08
; PRIOR APPLICATION NUMBER: 09/290,049
; PRIOR FILING DATE: 1999-04-12
; PRIOR APPLICATION NUMBER: 60/081,550
; PRIOR FILING DATE: 1998-04-13
; PRIOR APPLICATION NUMBER: 60/115,142
; PRIOR FILING DATE: 1999-01-08
; NUMBER OF SEQ ID NOS: 45
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 29

```

; LENGTH: 431
; TYPE: PRT
; ORGANISM: Streptococcus mutans
US-10-797-821-29

Query Match      100.0%; Score 94; DB 5; Length 431;
Best Local Similarity 100.0%; Pred. No. 1.3e-06;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KSNAAATSYINAIINSKVSVD 20
Db 113 KSNAAATSYINAIINSKVSVD 132

RESULT 7
US-10-797-821-30
; Sequence 30, Application US/10797821
; Publication No. US20050031633A1
; GENERAL INFORMATION:
; APPLICANT: Smith, Daniel J.
; APPLICANT: Taubman, Martin A.
; TITLE OF INVENTION: Glucan Binding Protein and Glycosyltransferase Immunogens
; FILE REFERENCE: 25669-020
; CURRENT APPLICATION NUMBER: US/10/797,821
; CURRENT FILING DATE: 2004-03-09
; PRIOR APPLICATION NUMBER: 10/383,930
; PRIOR FILING DATE: 2003-03-07
; PRIOR APPLICATION NUMBER: 60/363,209
; PRIOR FILING DATE: 2002-03-07
; PRIOR APPLICATION NUMBER: 60/402,483
; PRIOR FILING DATE: 2002-08-08
; PRIOR APPLICATION NUMBER: 09/290,049
; PRIOR FILING DATE: 1999-04-12
; PRIOR APPLICATION NUMBER: 60/081,550
; PRIOR FILING DATE: 1998-04-13
; PRIOR APPLICATION NUMBER: 60/115,142
; PRIOR FILING DATE: 1999-01-08
; NUMBER OF SEQ ID NOS: 45
; SOFTWARE: Patentin version 3.2
; SEQ ID NO 30
; LENGTH: 431
; TYPE: PRT
; ORGANISM: Streptococcus mutans
US-10-797-821-30

Query Match      100.0%; Score 94; DB 5; Length 431;
Best Local Similarity 100.0%; Pred. No. 1.3e-06;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KSNAAATSYINAIINSKVSVD 20
Db 113 KSNAAATSYINAIINSKVSVD 132

RESULT 8
US-10-797-821-33
; Sequence 33, Application US/10797821
; Publication No. US20050031633A1
; GENERAL INFORMATION:
; APPLICANT: Smith, Daniel J.
; APPLICANT: Taubman, Martin A.
; TITLE OF INVENTION: Glucan Binding Protein and Glycosyltransferase Immunogens
; FILE REFERENCE: 25669-020
; CURRENT APPLICATION NUMBER: US/10/797,821
; CURRENT FILING DATE: 2004-03-09
; PRIOR APPLICATION NUMBER: 10/383,930
; PRIOR FILING DATE: 2003-03-07
; PRIOR APPLICATION NUMBER: 60/363,209
; PRIOR FILING DATE: 2002-03-07
; PRIOR APPLICATION NUMBER: 60/402,483
; PRIOR FILING DATE: 2002-08-08
; PRIOR APPLICATION NUMBER: 09/290,049
; PRIOR FILING DATE: 1999-04-12

```

```

; PRIOR APPLICATION NUMBER: 60/081,550
; PRIOR FILING DATE: 1998-04-13
; ORGANISM: Streptococcus mutans
US-10-797-821-33

Query Match      100.0%; Score 94; DB 5; Length 431;
Best Local Similarity 100.0%; Pred. No. 1.3e-06;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KSNAAATSYINAIINSKVSVD 20
Db 113 KSNAAATSYINAIINSKVSVD 132

RESULT 9
US-10-383-930-31
; Sequence 31, Application US/10383930
; Publication No. US20040127400A1
; GENERAL INFORMATION:
; APPLICANT: Smith, Daniel J
; APPLICANT: Taubman, Martin A
; TITLE OF INVENTION: Immunogenicity of Glucan Binding Protein
; FILE REFERENCE: 25669-018
; CURRENT APPLICATION NUMBER: US/10/383,930
; CURRENT FILING DATE: 2003-03-07
; PRIOR APPLICATION NUMBER: 60/402,483
; PRIOR FILING DATE: 2002-08-08
; PRIOR APPLICATION NUMBER: 60/363,209
; PRIOR FILING DATE: 2002-03-07
; NUMBER OF SEQ ID NOS: 41
; SOFTWARE: Patentin version 3.2
; SEQ ID NO 31
; LENGTH: 432
; TYPE: PRT
; ORGANISM: Streptococcus mutans
US-10-383-930-31

Query Match      100.0%; Score 94; DB 4; Length 432;
Best Local Similarity 100.0%; Pred. No. 1.4e-06;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KSNAAATSYINAIINSKVSVD 20
Db 113 KSNAAATSYINAIINSKVSVD 132

RESULT 10
US-10-383-930-32
; Sequence 32, Application US/10383930
; Publication No. US20040127400A1
; GENERAL INFORMATION:
; APPLICANT: Smith, Daniel J
; APPLICANT: Taubman, Martin A
; TITLE OF INVENTION: Immunogenicity of Glucan Binding Protein
; FILE REFERENCE: 25669-018
; CURRENT APPLICATION NUMBER: US/10/383,930
; CURRENT FILING DATE: 2003-03-07
; PRIOR APPLICATION NUMBER: 60/402,483
; PRIOR FILING DATE: 2002-08-08
; PRIOR APPLICATION NUMBER: 60/363,209
; PRIOR FILING DATE: 2002-03-07
; NUMBER OF SEQ ID NOS: 41
; SOFTWARE: Patentin version 3.2
; SEQ ID NO 32
; LENGTH: 432
; TYPE: PRT

```

```

; ORGANISM: Streptococcus mutans
US-10-383-930-32

Query Match      100.0%; Score 94; DB 4; Length 432;
Best Local Similarity 100.0%; Pred. No. 1.4e-06;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KSNAAATSYINAIINSKVSVD 20
DB 113 KSNAAATSYINAIINSKVSVD 132

RESULT 11
US-10-797-821-31
; Sequence 31, Application US/10797821
; Publication No. US20050031633A1
; GENERAL INFORMATION:
; APPLICANT: Smith, Daniel J.
; APPLICANT: Taubman, Martin A.
; TITLE OF INVENTION: Glucan Binding Protein and Glycosyltransferase Immunogens
; FILE REFERENCE: 25669-020
; CURRENT APPLICATION NUMBER: US/10/797,821
; CURRENT FILING DATE: 2004-03-09
; PRIOR APPLICATION NUMBER: 10/383,930
; PRIOR FILING DATE: 2003-03-07
; PRIOR APPLICATION NUMBER: 60/363,209
; PRIOR FILING DATE: 2002-03-07
; PRIOR APPLICATION NUMBER: 60/402,483
; PRIOR FILING DATE: 2002-08-08
; PRIOR APPLICATION NUMBER: 09/290,049
; PRIOR FILING DATE: 1999-04-12
; PRIOR APPLICATION NUMBER: 60/081,550
; PRIOR FILING DATE: 1999-01-08
; NUMBER OF SEQ ID NOS: 45
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 31
; TYPE: PRT
; ORGANISM: Streptococcus mutans
US-10-797-821-31

Query Match      100.0%; Score 94; DB 5; Length 432;
Best Local Similarity 100.0%; Pred. No. 1.4e-06;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KSNAAATSYINAIINSKVSVD 20
DB 113 KSNAAATSYINAIINSKVSVD 132

RESULT 12
US-10-797-821-32
; Sequence 32, Application US/10797821
; Publication No. US20050031633A1
; GENERAL INFORMATION:
; APPLICANT: Smith, Daniel J.
; APPLICANT: Taubman, Martin A.
; TITLE OF INVENTION: Glucan Binding Protein and Glycosyltransferase Immunogens
; FILE REFERENCE: 25669-020
; CURRENT APPLICATION NUMBER: US/10/797,821
; CURRENT FILING DATE: 2004-03-09
; PRIOR APPLICATION NUMBER: 10/383,930
; PRIOR FILING DATE: 2003-03-07
; PRIOR APPLICATION NUMBER: 60/363,209
; PRIOR FILING DATE: 2002-03-07
; PRIOR APPLICATION NUMBER: 60/402,483
; PRIOR FILING DATE: 2002-08-08
; PRIOR APPLICATION NUMBER: 09/290,049
; PRIOR FILING DATE: 1999-04-12
; PRIOR APPLICATION NUMBER: 60/081,550
; PRIOR FILING DATE: 1998-04-13

```

```

; PRIOR APPLICATION NUMBER: 60/115,142
; PRIOR FILING DATE: 1999-01-08
; NUMBER OF SEQ ID NOS: 45
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 32
; LENGTH: 432
; TYPE: PRT
; ORGANISM: Streptococcus mutans
US-10-797-821-32

Query Match      100.0%; Score 94; DB 5; Length 432;
Best Local Similarity 100.0%; Pred. No. 1.4e-06;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KSNAAATSYINAIINSKVSVD 20
DB 113 KSNAAATSYINAIINSKVSVD 132

RESULT 13
US-10-383-930-11
; Sequence 11, Application US/10383930
; Publication No. US20040127400A1
; GENERAL INFORMATION:
; APPLICANT: Smith, Daniel J.
; APPLICANT: Taubman, Martin A.
; TITLE OF INVENTION: Immunogenicity of Glucan Binding Protein
; FILE REFERENCE: 25669-018
; CURRENT APPLICATION NUMBER: US/10/383,930
; CURRENT FILING DATE: 2003-03-07
; PRIOR APPLICATION NUMBER: 60/402,483
; PRIOR FILING DATE: 2002-08-08
; PRIOR APPLICATION NUMBER: 60/363,209
; PRIOR FILING DATE: 2002-03-07
; NUMBER OF SEQ ID NOS: 41
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 11
; LENGTH: 20
; TYPE: PRT
; ORGANISM: Streptococcus mutans
US-10-383-930-11

Query Match      79.8%; Score 75; DB 4; Length 20;
Best Local Similarity 100.0%; Pred. No. 4.8e-05;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 5 ATSYINAIINSKVSVD 20
DB 1 ATSYINAIINSKVSVD 16

RESULT 14
US-10-797-821-11
; Sequence 11, Application US/10797821
; Publication No. US20050031633A1
; GENERAL INFORMATION:
; APPLICANT: Smith, Daniel J.
; APPLICANT: Taubman, Martin A.
; TITLE OF INVENTION: Glucan Binding Protein and Glycosyltransferase Immunogens
; FILE REFERENCE: 25669-020
; CURRENT APPLICATION NUMBER: US/10/797,821
; CURRENT FILING DATE: 2004-03-09
; PRIOR APPLICATION NUMBER: 10/383,930
; PRIOR FILING DATE: 2003-03-07
; PRIOR APPLICATION NUMBER: 60/363,209
; PRIOR FILING DATE: 2002-03-07
; PRIOR APPLICATION NUMBER: 60/402,483
; PRIOR FILING DATE: 2002-08-08
; PRIOR APPLICATION NUMBER: 09/290,049
; PRIOR FILING DATE: 1999-04-12
; PRIOR APPLICATION NUMBER: 60/081,550
; PRIOR FILING DATE: 1998-04-13

```


; PRIOR FILING DATE: 1999-01-08
 ; NUMBER OF SEQ ID NOS: 45
 ; SOFTWARE: PatentIn version 3.2
 ; SEQ ID NO 11
 ; LENGTH: 20
 ; TYPE: PRT
 ; ORGANISM: Artificial
 ; FEATURE:
 ; OTHER INFORMATION: GbpB peptide
 US-10-797-821-11

Query Match 79.8%; Score 75; DB 5; Length 20;
 Best Local Similarity 100.0%; Pred. No. 4.8e-05;
 Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 5 ATSYINAINSKVSD 20
 DB 1 ATSYINAINSKVSD 16

RESULT 15
 US-10-474-792-600
 ; Sequence 600, Application US/10474792
 ; Publication No. US20040236072A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Olmsted, Stephen
 ; APPLICANT: Zagursky, Robert
 ; APPLICANT: Nickbarg, Elliot
 ; APPLICANT: Winter, Lourie
 ; TITLE OF INVENTION: SURFACE PROTEINS OF STREPTOCOCCUS PYOGENES
 ; FILE REFERENCE: AM 100399
 ; CURRENT APPLICATION NUMBER: US/10/474,792
 ; CURRENT FILING DATE: 2003-10-14
 ; NUMBER OF SEQ ID NOS: 674
 ; SOFTWARE: PatentIn version 3.0
 ; SEQ ID NO 600
 ; LENGTH: 398
 ; TYPE: PRT
 ; ORGANISM: Streptococcus pyogenes
 US-10-474-792-600

Query Match 73.4%; Score 69; DB 5; Length 398;
 Best Local Similarity 60.0%; Pred. No. 0.013;
 Matches 12; Conservative 5; Mismatches 3; Indels 0; Gaps 0;

QY 1 KSNAAATSYINAINSKVSD 20
 DB 110 KNETSGYINALLNSKSID 129

Search completed: January 25, 2006, 20:10:00
 Job time : 60 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2006 Compugen Ltd.

OM protein - protein search, using sw model

Run on: January 25, 2006, 19:15:26 ; Search time 5.5 Seconds
(without alignments)
39.378 Million cell updates/sec

Title: US-10-797-821-1
Perfect score: 94
Sequence: 1 KSNATSYINAINSKSVSD 20

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 75621 seqs, 10829074 residues

Total number of hits satisfying chosen parameters: 75621

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published Applications AA New.*
1: /cgn2_6/ptodata/2/pubpaa/US08_NEW_PUB_PEP.*
2: /cgn2_6/ptodata/2/pubpaa/US06_NEW_PUB_PEP.*
3: /cgn2_6/ptodata/2/pubpaa/US07_NEW_PUB_PEP.*
4: /cgn2_6/ptodata/2/pubpaa/US09_NEW_PUB_PEP.*
5: /cgn2_6/ptodata/2/pubpaa/US10_NEW_PUB_PEP.*
6: /cgn2_6/ptodata/2/pubpaa/US11_NEW_PUB_PEP.*
7: /cgn2_6/ptodata/2/pubpaa/US12_NEW_PUB_PEP.*
8: /cgn2_6/ptodata/2/pubpaa/US13_NEW_PUB_PEP.*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	94	100.0	431	7	US-11-052-554A-210
2	69	73.4	398	7	US-11-052-554A-210
3	38	40.4	91	7	US-11-052-554A-252
4	38	40.4	124	7	US-11-212-443-163
5	38	40.4	124	7	US-11-212-443-14
6	38	40.4	186	6	US-11-212-443-16
7	38	40.4	273	7	US-10-467-657-470
8	38	40.4	342	7	US-11-196-475-7
9	38	40.4	434	7	US-11-196-475-170
10	38	40.4	446	7	US-11-196-475-150
11	38	40.4	450	7	US-11-196-475-148
12	38	40.4	453	7	US-11-196-475-146
13	37	39.4	702	7	US-11-033-039-282
14	37	39.4	706	6	US-10-485-517-146
15	37	39.4	908	6	US-10-821-234-1267
16	37	39.4	908	6	US-10-793-626-1304
17	36	38.3	325	7	US-11-132-864-31
18	36	38.3	348	6	US-10-793-626-1702
19	36	38.3	351	6	US-10-793-626-338
20	36	38.3	353	7	US-11-132-864-29
21	36	38.3	330	6	US-10-131-826A-338
22	36	38.3	637	7	US-11-175-690-266
23	36	38.3	751	7	US-11-052-554A-109
24	36	38.3	195	6	US-10-873-528-139
25	35.5	37.8			

26	35.5	37.8	252	6	US-10-401-386B-38	Sequence 38, Appl
27	35.5	37.8	255	6	US-10-401-386B-34	Sequence 34, Appl
28	35	37.2	41	6	US-10-916-827-27	Sequence 27, Appl
29	35	37.2	41	6	US-10-916-827-45	Sequence 45, Appl
30	35	37.2	65	6	US-10-667-295-7	Sequence 7, Appl
31	35	37.2	120	7	US-11-054-669-114	Sequence 114, Appl
32	35	37.2	286	7	US-11-089-551A-27	Sequence 27, Appl
33	35	37.2	329	7	US-11-156-084-28	Sequence 28, Appl
34	35	37.2	329	7	US-11-156-084-46	Sequence 46, Appl
35	35	37.2	329	7	US-11-156-084-171	Sequence 171, Appl
36	35	37.2	525	7	US-11-082-389-36	Sequence 36, Appl
37	35	37.2	543	6	US-10-495-664-3	Sequence 3, Appl
38	35	37.2	545	7	US-11-082-389-34	Sequence 34, Appl
39	35	37.2	693	6	US-10-873-528-185	Sequence 185, Appl
40	35	37.2	786	6	US-10-467-962B-103	Sequence 103, Appl
41	35	37.2	1107	6	US-10-485-517-145	Sequence 145, Appl
42	35	37.2	1185	6	US-10-877-346-7	Sequence 7, Appl
43	35	37.2	2080	6	US-10-821-234-1640	Sequence 1640, Ap
44	34.5	36.7	446	6	US-10-793-626-2250	Sequence 2250, Ap
45	34.5	36.7	643	7	US-11-052-554A-106	Sequence 106, App

ALIGNMENTS

RESULT 1
US-11-052-554A-210
; Sequence 210, Application US/11052554A
; Publication No. US2005028866A1
; GENERAL INFORMATION:
; APPLICANT: Sachdeva, et al.
; TITLE OF INVENTION: COMPUTATIONAL METHOD FOR IDENTIFYING ADHESIN AND ADHESIN-LIKE
; FILE REFERENCE: 30853/40359A
; CURRENT APPLICATION NUMBER: US/11/052,554A
; CURRENT FILING DATE: 2005-02-07
; PRIOR APPLICATION NUMBER: US 60/589,227
; PRIOR FILING DATE: 2004-07-20
; PRIOR APPLICATION NUMBER: IN 173/DEL/2004
; PRIOR FILING DATE: 2004-02-06
; NUMBER OF SEQ ID NOS: 763
; SOFTWARE: Patentn version 3.3
; SEQ ID NO 210
; LENGTH: 431
; TYPE: PRT
; ORGANISM: Streptococcus mutans UA159
US-11-052-554A-210

Query Match 100.0%; Score 94; DB 7; Length 431;
Best Local Similarity 100.0%; Pred. No. 2.8e-08;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KSNATSYINAINSKSVSD 20
Db 113 KSNATSYINAINSKSVSD 132

RESULT 2
US-11-052-554A-252
; Sequence 252, Application US/11052554A
; Publication No. US2005028866A1
; GENERAL INFORMATION:
; APPLICANT: Sachdeva, et al.
; TITLE OF INVENTION: COMPUTATIONAL METHOD FOR IDENTIFYING ADHESIN AND ADHESIN-LIKE
; FILE REFERENCE: 30853/40359A
; CURRENT APPLICATION NUMBER: US/11/052,554A
; CURRENT FILING DATE: 2005-02-07
; PRIOR APPLICATION NUMBER: US 60/589,227
; PRIOR FILING DATE: 2004-07-20
; PRIOR APPLICATION NUMBER: IN 173/DEL/2004
; PRIOR FILING DATE: 2004-02-06
; NUMBER OF SEQ ID NOS: 763

✓

```

Query Match      40.4%; Score 38; DB 6; Length 186;
Best Local Similarity 53.8%; Pred. No. 16;
Matches 7; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

QY      6 TSYINAIINIKSV 18
      |:|:|:|:|
Db      118 TTYASAVNGKRV 130

RESULT 7
US-11-196-475-7
; Sequence 7, Application US/11196475
; Publication No. US20050271682A1
; GENERAL INFORMATION:
; APPLICANT: Dattwyler, Raymond J.
; APPLICANT: Gomes Solecki, Maria J. C.
; APPLICANT: Luft, Benjamin J.
; APPLICANT: Dunn, John J.
; TITLE OF INVENTION: Recombinant Constructs of Borrelia
; FILE REFERENCE: 2631.1001-011
; CURRENT APPLICATION NUMBER: US/11/196,475
; CURRENT FILING DATE: 2005-08-03
; PRIOR APPLICATION NUMBER: US 08/148,191
; PRIOR FILING DATE: 1993-11-01
; PRIOR APPLICATION NUMBER: US 08/235,836
; PRIOR FILING DATE: 1994-04-29
; PRIOR APPLICATION NUMBER: US 09/666,017
; PRIOR FILING DATE: 2000-09-19
; PRIOR APPLICATION NUMBER: US 60/226,484
; PRIOR FILING DATE: 2000-08-18
; PRIOR APPLICATION NUMBER: PCT/US01/24736
; PRIOR FILING DATE: 2001-08-07
; NUMBER OF SEQ ID NOS: 213
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 7
; LENGTH: 273
; TYPE: PRT
; ORGANISM: Borrelia burgdorferi
US-11-196-475-7

Query Match      40.4%; Score 38; DB 7; Length 273;
Best Local Similarity 38.9%; Pred. No. 26;
Matches 7; Conservative 3; Mismatches 8; Indels 0; Gaps 0;

QY      3 NAATSYINAIINIKSVSD 20
      |:|:|:|:|
Db      217 NSGTSTLTITVNSKTKD 234

RESULT 8
US-11-196-475-170
; Sequence 170, Application US/11196475
; Publication No. US20050271682A1
; GENERAL INFORMATION:
; APPLICANT: Dattwyler, Raymond J.
; APPLICANT: Gomes Solecki, Maria J. C.
; APPLICANT: Luft, Benjamin J.
; APPLICANT: Dunn, John J.
; TITLE OF INVENTION: Recombinant Constructs of Borrelia
; FILE REFERENCE: 2631.1001-011
; CURRENT APPLICATION NUMBER: US/11/196,475
; CURRENT FILING DATE: 2005-08-03
; PRIOR APPLICATION NUMBER: US 08/148,191
; PRIOR FILING DATE: 1993-11-01
; PRIOR APPLICATION NUMBER: US 08/235,836
; PRIOR FILING DATE: 1994-04-29
; PRIOR APPLICATION NUMBER: US 09/666,017
; PRIOR FILING DATE: 2000-09-19
; PRIOR APPLICATION NUMBER: US 60/226,484
; PRIOR FILING DATE: 2000-08-18
; PRIOR APPLICATION NUMBER: PCT/US01/24736
; PRIOR FILING DATE: 2001-08-07
; NUMBER OF SEQ ID NOS: 213
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 7
; LENGTH: 273
; TYPE: PRT
; ORGANISM: Borrelia burgdorferi
US-11-196-475-170

Query Match      40.4%; Score 38; DB 7; Length 434;
Best Local Similarity 38.9%; Pred. No. 45;
Matches 7; Conservative 3; Mismatches 8; Indels 0; Gaps 0;

QY      3 NAATSYINAIINIKSVSD 20
      |:|:|:|:|
Db      378 NSGTSTLTITVNSKTKD 395

RESULT 10
US-11-196-475-150
; Sequence 150, Application US/11196475
; Publication No. US20050271682A1
; GENERAL INFORMATION:
; APPLICANT: Dattwyler, Raymond J.
; APPLICANT: Gomes Solecki, Maria J. C.
; APPLICANT: Luft, Benjamin J.
; APPLICANT: Dunn, John J.

```

```

Query Match      40.4%; Score 38; DB 6; Length 186;
Best Local Similarity 53.8%; Pred. No. 16;
Matches 7; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

QY      6 TSYINAIINIKSV 18
      |:|:|:|:|
Db      118 TTYASAVNGKRV 130

RESULT 7
US-11-196-475-7
; Sequence 7, Application US/11196475
; Publication No. US20050271682A1
; GENERAL INFORMATION:
; APPLICANT: Dattwyler, Raymond J.
; APPLICANT: Gomes Solecki, Maria J. C.
; APPLICANT: Luft, Benjamin J.
; APPLICANT: Dunn, John J.
; TITLE OF INVENTION: Recombinant Constructs of Borrelia
; FILE REFERENCE: 2631.1001-011
; CURRENT APPLICATION NUMBER: US/11/196,475
; CURRENT FILING DATE: 2005-08-03
; PRIOR APPLICATION NUMBER: US 08/148,191
; PRIOR FILING DATE: 1993-11-01
; PRIOR APPLICATION NUMBER: US 08/235,836
; PRIOR FILING DATE: 1994-04-29
; PRIOR APPLICATION NUMBER: US 09/666,017
; PRIOR FILING DATE: 2000-09-19
; PRIOR APPLICATION NUMBER: US 60/226,484
; PRIOR FILING DATE: 2000-08-18
; PRIOR APPLICATION NUMBER: PCT/US01/24736
; PRIOR FILING DATE: 2001-08-07
; NUMBER OF SEQ ID NOS: 213
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 7
; LENGTH: 273
; TYPE: PRT
; ORGANISM: Borrelia burgdorferi
US-11-196-475-7

Query Match      40.4%; Score 38; DB 7; Length 273;
Best Local Similarity 38.9%; Pred. No. 26;
Matches 7; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

QY      3 NAATSYINAIINIKSVSD 20
      |:|:|:|:|
Db      217 NSGTSTLTITVNSKTKD 234

RESULT 8
US-11-196-475-170
; Sequence 170, Application US/11196475
; Publication No. US20050271682A1
; GENERAL INFORMATION:
; APPLICANT: Dattwyler, Raymond J.
; APPLICANT: Gomes Solecki, Maria J. C.
; APPLICANT: Luft, Benjamin J.
; APPLICANT: Dunn, John J.
; TITLE OF INVENTION: Recombinant Constructs of Borrelia
; FILE REFERENCE: 2631.1001-011
; CURRENT APPLICATION NUMBER: US/11/196,475
; CURRENT FILING DATE: 2005-08-03
; PRIOR APPLICATION NUMBER: US 08/148,191
; PRIOR FILING DATE: 1993-11-01
; PRIOR APPLICATION NUMBER: US 08/235,836
; PRIOR FILING DATE: 1994-04-29
; PRIOR APPLICATION NUMBER: US 09/666,017
; PRIOR FILING DATE: 2000-09-19
; PRIOR APPLICATION NUMBER: US 60/226,484
; PRIOR FILING DATE: 2000-08-18
; PRIOR APPLICATION NUMBER: PCT/US01/24736
; PRIOR FILING DATE: 2001-08-07
; NUMBER OF SEQ ID NOS: 213
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 7
; LENGTH: 273
; TYPE: PRT
; ORGANISM: Borrelia burgdorferi
US-11-196-475-170

Query Match      40.4%; Score 38; DB 6; Length 186;
Best Local Similarity 53.8%; Pred. No. 16;
Matches 7; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

QY      6 TSYINAIINIKSV 18
      |:|:|:|:|
Db      118 TTYASAVNGKRV 130

RESULT 7
US-11-196-475-7
; Sequence 7, Application US/11196475
; Publication No. US20050271682A1
; GENERAL INFORMATION:
; APPLICANT: Dattwyler, Raymond J.
; APPLICANT: Gomes Solecki, Maria J. C.
; APPLICANT: Luft, Benjamin J.
; APPLICANT: Dunn, John J.
; TITLE OF INVENTION: Recombinant Constructs of Borrelia
; FILE REFERENCE: 2631.1001-011
; CURRENT APPLICATION NUMBER: US/11/196,475
; CURRENT FILING DATE: 2005-08-03
; PRIOR APPLICATION NUMBER: US 08/148,191
; PRIOR FILING DATE: 1993-11-01
; PRIOR APPLICATION NUMBER: US 08/235,836
; PRIOR FILING DATE: 1994-04-29
; PRIOR APPLICATION NUMBER: US 09/666,017
; PRIOR FILING DATE: 2000-09-19
; PRIOR APPLICATION NUMBER: US 60/226,484
; PRIOR FILING DATE: 2000-08-18
; PRIOR APPLICATION NUMBER: PCT/US01/24736
; PRIOR FILING DATE: 2001-08-07
; NUMBER OF SEQ ID NOS: 213
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 7
; LENGTH: 273
; TYPE: PRT
; ORGANISM: Borrelia burgdorferi
US-11-196-475-7

Query Match      40.4%; Score 38; DB 7; Length 434;
Best Local Similarity 38.9%; Pred. No. 45;
Matches 7; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

QY      3 NAATSYINAIINIKSVSD 20
      |:|:|:|:|
Db      378 NSGTSTLTITVNSKTKD 395

RESULT 10
US-11-196-475-150
; Sequence 150, Application US/11196475
; Publication No. US20050271682A1
; GENERAL INFORMATION:
; APPLICANT: Dattwyler, Raymond J.
; APPLICANT: Gomes Solecki, Maria J. C.
; APPLICANT: Luft, Benjamin J.
; APPLICANT: Dunn, John J.

```

```

; TITLE OF INVENTION: Recombinant Constructs of Borrelia
; TITLE OF INVENTION: Burgdorferi
; FILE REFERENCE: 2631.1001-011
; CURRENT APPLICATION NUMBER: US/11/196,475
; CURRENT FILING DATE: 2005-08-03
; PRIOR APPLICATION NUMBER: US 08/148,191
; PRIOR FILING DATE: 1993-11-01
; PRIOR APPLICATION NUMBER: US 08/235,836
; PRIOR FILING DATE: 1994-04-29
; PRIOR APPLICATION NUMBER: US 09/666,017
; PRIOR FILING DATE: 2000-09-19
; PRIOR APPLICATION NUMBER: US 60/226,484
; PRIOR FILING DATE: 2000-08-18
; PRIOR APPLICATION NUMBER: PCT/US01/24736
; PRIOR FILING DATE: 2001-08-07
; NUMBER OF SEQ ID NOS: 213
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 150
; LENGTH: 446
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Chimeric Protein
US-11-196-475-150

Query Match 40.4%; Score 38; DB 7; Length 446;
Best Local Similarity 38.9%; Pred. No. 46;
Matches 7; Conservative 3; Mismatches 8; Indels 0; Gaps 0;

QY 3 NAATSYINAINSKSVSD 20
| : || : : ||| |
DB 390 NSGTSTLTITVNSKTKD 407

RESULT 11
US-11-196-475-148
; Sequence 148, Application US/11/196475
; Publication No. US20050271682A1
; GENERAL INFORMATION:
; APPLICANT: Dattwyler, Raymond J.
; APPLICANT: Gomes Solecki, Maria J. C.
; APPLICANT: Luft, Benjamin J.
; APPLICANT: Dunn, John J.
; TITLE OF INVENTION: Recombinant Constructs of Borrelia
; TITLE OF INVENTION: Burgdorferi
; FILE REFERENCE: 2631.1001-011
; CURRENT APPLICATION NUMBER: US/11/196,475
; CURRENT FILING DATE: 2005-08-03
; PRIOR APPLICATION NUMBER: US 08/148,191
; PRIOR FILING DATE: 1993-11-01
; PRIOR APPLICATION NUMBER: US 08/235,836
; PRIOR FILING DATE: 1994-04-29
; PRIOR APPLICATION NUMBER: US 09/666,017
; PRIOR FILING DATE: 2000-09-19
; PRIOR APPLICATION NUMBER: US 60/226,484
; PRIOR FILING DATE: 2000-08-18
; PRIOR APPLICATION NUMBER: PCT/US01/24736
; PRIOR FILING DATE: 2001-08-07
; NUMBER OF SEQ ID NOS: 213
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 148
; LENGTH: 450
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Chimeric Protein
US-11-196-475-148

Query Match 40.4%; Score 38; DB 7; Length 450;
Best Local Similarity 38.9%; Pred. No. 46;
Matches 7; Conservative 3; Mismatches 8; Indels 0; Gaps 0;

QY 3 NAATSYINAINSKSVSD 20
| : || : : ||| |
DB 390 NSGTSTLTITVNSKTKD 407

```

```

| : || : : ||| |
DB 394 NSGTSTLTITVNSKTKD 411

RESULT 12
US-11-196-475-146
; Sequence 146, Application US/11/196475
; Publication No. US20050271682A1
; GENERAL INFORMATION:
; APPLICANT: Dattwyler, Raymond J.
; APPLICANT: Gomes Solecki, Maria J. C.
; APPLICANT: Luft, Benjamin J.
; APPLICANT: Dunn, John J.
; TITLE OF INVENTION: Recombinant Constructs of Borrelia
; TITLE OF INVENTION: Burgdorferi
; FILE REFERENCE: 2631.1001-011
; CURRENT APPLICATION NUMBER: US/11/196,475
; CURRENT FILING DATE: 2005-08-03
; PRIOR APPLICATION NUMBER: US 08/148,191
; PRIOR FILING DATE: 1993-11-01
; PRIOR APPLICATION NUMBER: US 08/235,836
; PRIOR FILING DATE: 1994-04-29
; PRIOR APPLICATION NUMBER: US 09/666,017
; PRIOR FILING DATE: 2000-09-19
; PRIOR APPLICATION NUMBER: US 60/226,484
; PRIOR FILING DATE: 2000-08-18
; PRIOR APPLICATION NUMBER: PCT/US01/24736
; PRIOR FILING DATE: 2001-08-07
; NUMBER OF SEQ ID NOS: 213
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 146
; LENGTH: 453
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Chimeric Protein
US-11-196-475-146

Query Match 40.4%; Score 38; DB 7; Length 453;
Best Local Similarity 38.9%; Pred. No. 47;
Matches 7; Conservative 3; Mismatches 8; Indels 0; Gaps 0;

QY 3 NAATSYINAINSKSVSD 20
| : || : : ||| |
DB 397 NSGTSTLTITVNSKTKD 414

RESULT 13
US-11-196-475-112
; Sequence 112, Application US/11/196475
; Publication No. US20050271682A1
; GENERAL INFORMATION:
; APPLICANT: Dattwyler, Raymond J.
; APPLICANT: Gomes Solecki, Maria J. C.
; APPLICANT: Luft, Benjamin J.
; APPLICANT: Dunn, John J.
; TITLE OF INVENTION: Recombinant Constructs of Borrelia
; TITLE OF INVENTION: Burgdorferi
; FILE REFERENCE: 2631.1001-011
; CURRENT APPLICATION NUMBER: US/11/196,475
; CURRENT FILING DATE: 2005-08-03
; PRIOR APPLICATION NUMBER: US 08/148,191
; PRIOR FILING DATE: 1993-11-01
; PRIOR APPLICATION NUMBER: US 08/235,836
; PRIOR FILING DATE: 1994-04-29
; PRIOR APPLICATION NUMBER: US 09/666,017
; PRIOR FILING DATE: 2000-09-19
; PRIOR APPLICATION NUMBER: US 60/226,484
; PRIOR FILING DATE: 2000-08-18
; PRIOR APPLICATION NUMBER: PCT/US01/24736
; PRIOR FILING DATE: 2001-08-07
; NUMBER OF SEQ ID NOS: 213
; SOFTWARE: FastSeq for Windows Version 4.0

```

```
; SEQ ID NO 112
; LENGTH: 573
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Chimeric Protein
US-11-196-475-112

Query Match      40.4%; Score 38; DB 7; Length 573;
Best Local Similarity 38.9%; Pred. No. 62;
Matches 7; Conservative 3; Mismatches 8; Indels 0; Gaps 0;

QY      3 NAATSYINALINSKVSVD 20
Db      200 NSGTSTLTITVNSKTKD 217

RESULT 14
US-11-033-039-282
; Sequence 282, Application US/11033039
; Publication No. US20060002947A1
; GENERAL INFORMATION:
; APPLICANT: HUMPHREYS, ROBERT
; APPLICANT: XU, MINZHEN
; TITLE OF INVENTION: LI-KEY/ANTIGENIC EPITOPE HYBRID PEPTIDE VACCINES
; FILE REFERENCE: REH-2017US01
; CURRENT APPLICATION NUMBER: US/11/033,039
; CURRENT FILING DATE: 2005-01-11
; PRIOR APPLICATION NUMBER: 10/245,871
; PRIOR FILING DATE: 2002-09-17
; PRIOR APPLICATION NUMBER: 10/197,000
; PRIOR FILING DATE: 2002-07-17
; PRIOR APPLICATION NUMBER: 09/396,813
; PRIOR FILING DATE: 1999-09-14
; NUMBER OF SEQ ID NOS: 1452
; SOFTWARE: PatentIn version 3.3
;:SEQ ID NO 282
; LENGTH: 702
; TYPE: PRT
; ORGANISM: Homo sapiens
US-11-033-039-282

Query Match      39.4%; Score 37; DB 7; Length 702;
Best Local Similarity 50.0%; Pred. No. 1.le+02;
Matches 9; Conservative 3; Mismatches 6; Indels 0; Gaps 0;

QY      2 SNAATSYINALINSKVS 19
Db      658 SNLATGRNNSIVAKSITVS 675

RESULT 15
US-10-485-517-146
; Sequence 146, Application US/10485517
; Publication No. US20050256299A1
; GENERAL INFORMATION:
; APPLICANT: University of Sheffield
; APPLICANT: Biosynexus Incorporated
; APPLICANT: Foster, Simon
; APPLICANT: Mond, James
; TITLE OF INVENTION: Antigenic Polypeptides
; FILE REFERENCE: P100629WO
; CURRENT APPLICATION NUMBER: US/10/485,517
; CURRENT FILING DATE: 2004-02-02
; PRIOR APPLICATION NUMBER: GB 0118825.9
; PRIOR FILING DATE: 2001-08-02
; PRIOR APPLICATION NUMBER: GB 0200349.9
; PRIOR FILING DATE: 2002-01-09
; NUMBER OF SEQ ID NOS: 424
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 146
; LENGTH: 706
; TYPE: PRT
```

```
; ORGANISM: Staphylococcus aureus
US-10-485-517-146

Query Match      39.4%; Score 37; DB 6; Length 706;
Best Local Similarity 35.0%; Pred. No. 1.2e+02;
Matches 7; Conservative 6; Mismatches 7; Indels 0; Gaps 0;

QY      1 KSNATSYINALINSKVSVD 20
Db      347 QKNALTQQVNSAQNVQAVND 366

Search completed: January 25, 2006, 20:11:07
Job time : 6.5 secs
```

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2006 Compugen Ltd.

OM protein - protein search, using sw model

Run on: January 25, 2006, 18:48:24 ; Search time 68.2 Seconds
(without alignments)
128.850 Million cell updates/sec

Title: US-10-797-821-2
Perfect score: 97
Sequence: 1 KHLITIQQVSALQTOQAG 20

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 2443163 seqs, 439378781 residues

Total number of hits satisfying chosen parameters: 2443163

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : A_Geneseq_21.*
1: Geneseqp1980s:*
2: Geneseqp1990s:*
3: Geneseqp2000s:*
4: Geneseqp2001s:*
5: Geneseqp2002s:*
6: Geneseqp2003as:*
7: Geneseqp2003bs:*
8: Geneseqp2004s:*
9: Geneseqp2005s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	97	100.0	20	7	Add93622 Streptoco
2	97	100.0	20	9	Adx37245 Streptoco
3	65	67.0	20	7	Add93629 Streptoco
4	65	67.0	20	9	Adx37252 Streptoco
5	65	67.0	431	7	Add93649 Streptoco
6	65	67.0	431	7	Add93650 Streptoco
7	65	67.0	431	9	Add93653 Streptoco
8	65	67.0	431	9	Adx37272 Streptoco
9	65	67.0	431	9	Adx37273 Streptoco
10	65	67.0	431	9	Adx37276 Streptoco
11	65	67.0	431	9	Aeb91500 Microbial
12	65	67.0	432	7	Add93651 Streptoco
13	65	67.0	432	7	Add93652 Streptoco
14	65	67.0	432	9	Adx37274 Streptoco
15	65	67.0	432	9	Adx37275 Streptoco
16	51	52.6	20	7	Add93628 Streptoco
17	51	52.6	20	9	Adx37251 Streptoco
18	48	49.5	100	5	ABP03636 Human ORF
19	46	47.4	604	6	Aae36324 Arabidops
20	45	46.4	447	5	ABP29684 Streptoco
21	45	46.4	447	5	Adu69524 S agalact
22	45	46.4	447	8	Adv88392 Streptoco
23	45	46.4	447	8	Adv81808 Streptoco
24	45	46.4	447	8	Adv79645 Streptoco

25	45	46.4	1034	6	ABU47461	Abu47461 Protein e
26	45	46.4	1115	3	AA93947	Aay93947 Amino aci
27	44	45.4	295	7	ADH88734	Adh88734 Enterococ
28	44	45.4	410	2	AA116613	Aay16613 Amino aci
29	44	45.4	412	6	ABU39026	Abu39026 Protein e
30	44	45.4	456	4	ABB52485	Abb52485 Escherich
31	44	45.4	456	3	AB124448	Ab124448 Human hho
32	44	45.4	785	3	ABO99688	Aab099688 Human bra
33	44	45.4	789	4	AAU39059	Aam39059 Human pol
34	44	45.4	789	4	AAU28092	Aau28092 Novel hum
35	44	45.4	789	7	ADF69106	Adf69106 Human MP5
36	44	45.4	834	8	ADX93182	Adx93182 Plant ful
37	44	45.4	851	8	ADS24336	Ads24336 Bacterial
38	44	45.4	1526	7	ABM88482	Abm88482 Rice abio
39	43	44.3	395	5	ABP25918	Abp25918 Streptoco
40	43	44.3	872	5	ABG77915	Abg77915 High leve
41	43	44.3	872	8	ADM48066	Adm48066 Polypepti
42	43	44.3	872	8	ADN20109	Adn20109 Bacterial
43	42.5	43.8	418	8	ADX79199	Adx79199 Plant ful
44	42.5	43.8	700	7	ABO83701	Abo83701 Pseudomon
45	42.5	43.8	964	4	AAE03671	Aae03671 Human TRA

ALIGNMENTS

RESULT 1
ADD93622
ID ADD93622 standard; peptide; 20 AA.
XX
AC ADD93622;
XX
DT 29-JAN-2004 (first entry)
XX
DE Streptococcus mutans glucan binding protein-B peptide fragment.
XX
KW Glucan binding protein-B; GbpB; vaccine; anticaries; epitope; immunogen.
XX
OS Streptococcus mutans.
XX
PN WO2003075845-A2.
XX
PD 18-SEP-2003.
XX
PF 07-MAR-2003; 2003WO-US006962.
XX
PR 07-MAR-2002; 2002US-0363209P.
PR 08-AUG-2002; 2002US-0402483P.
XX
(FORS-) FORSYTH INST.
Smith DU, Taubman MA;
WPI; 2003-845091/78.
Composition useful as vaccines for dental caries comprises a fragment of a glucan binding protein-B binding to a major histocompatibility complex class II protein.
Claim 11; Page 10; 49pp; English.

The present sequence is that of a peptide derived from the glucan binding protein-B (GbpB) of Streptococcus mutans strain SK32 ADD93649. The peptide binds to a major histocompatibility complex (MHC) class II protein. It was identified as a potential B cell epitope using a matrix-based algorithm for epitope prediction, which was used to search the primary amino acid sequence of GbpB for known MHC class II binding motifs. The peptide can be used in immunogenic compositions and subunit vaccines for dental caries. These comprise MHC class II protein-binding GbpB peptides covalently linked with peptide subunits (preferably from the catalytic domain) of a glucosyltransferase. The compositions are used in a claimed method of eliciting production of an antibody in a mammal. Diepitopic or multiepitopic polypeptides can be prepared synthetically or

CC by recombinant DNA technology. Antibodies raised against MHC class II
 CC binding fragments of GbpB can be used in passive immunisation.

XX SQ Sequence 20 AA;
 Query Match 100.0%; Score 97; DB 7; Length 20;
 Best Local Similarity 100.0%; Pred. No. 6e-09;
 Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 KHKLITIQGVSAALQTQQA 20
 |||||
 Db 1 KHKLITIQGVSAALQTQQA 20
 |||||

RESULT 2
 ADX37245
 ID ADX37245 standard; peptide; 20 AA.

XX AC ADX37245;

XX DT 21-APR-2005 (first entry)

XX DE Streptococcus mutant glucan binding protein B peptide #2.

XX KW immunogenicity; immune stimulation; glucan binding protein-B;
 KW microparticle; major histocompatibility complex; tooth disease.

XX OS Streptococcus mutans.

XX PN US2005031633-A1.

XX PD 10-FEB-2005.

XX PF 09-MAR-2004; 2004US-00797821.

XX PR 13-APR-1998; 98US-0081550P.

XX PR 08-JAN-1999; 99US-0115142P.

XX PR 12-APR-1999; 99US-00290049.

XX PR 07-MAR-2002; 2002US-0363209P.

XX PR 08-AUG-2002; 2002US-0402483P.

XX PR 07-MAR-2003; 2003US-00383930.

XX PA (SMIT/) SMITH D J.

XX PA (TAUB/) TAUBMAN M A.

XX PI Smith DJ, Taubman MA;

XX PI WPI; 2005-151644/16.

XX PS Claim 4; SEQ ID NO 2; 73pp; English.

XX CC The invention relates to a composition comprising a fragment of a glucan
 CC binding protein-B (GbpB) and a biocompatible microparticle, where the
 CC fragment binds to a major histocompatibility complex (MHC) class II
 CC protein. The composition is useful for producing an antibody for
 CC immunizing mammals against dental caries. This sequence corresponds to a
 CC Streptococcus mutans GbpB-derived peptide of the invention.

XX SQ Sequence 20 AA;

Query Match 100.0%; Score 97; DB 9; Length 20;
 Best Local Similarity 100.0%; Pred. No. 6e-09;
 Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 KHKLITIQGVSAALQTQQA 20
 |||||

Db 1 KHKLITIQGVSAALQTQQA 20
 |||||

RESULT 3

ADD93629
 ID ADD93629 standard; peptide; 20 AA.

XX AC ADD93629;

XX DT 29-JAN-2004 (first entry)

XX DE Streptococcus mutans glucan binding protein-B peptide fragment.

XX KW Glucan binding protein-B; GbpB; vaccine; anticaries; epitope; immunogen.

XX OS Streptococcus mutans.

XX PN WO2003075845-A2.

XX PD 18-SEP-2003.

XX PF 07-MAR-2003; 2003WO-US006962.

XX PR 07-MAR-2002; 2002US-0363209P.

XX PR 08-AUG-2002; 2002US-0402483P.

XX PA (FORS-) FORSYTH INST.

XX PI Smith DJ, Taubman MA;

XX PI WPI; 2003-845091/78.

XX PT Composition useful as vaccines for dental caries comprises a fragment of
 PT a glucan binding protein-B binding to a major histocompatibility complex
 PT class II protein.

XX PS Claim 11; Page 10; 49pp; English.

XX CC The present sequence is that of a peptide comprising amino acid residues
 CC 52-71 of the glucan binding protein-B (GbpB) of Streptococcus mutans
 CC strain SK32 ADD93649. The peptide binds to a major histocompatibility
 CC complex (MHC) class II protein. It was identified as a potential B cell
 CC epitope using a matrix-based algorithm for epitope prediction, which was
 CC used to search the primary amino acid sequence of GbpB for known MHC
 CC class II binding motifs. The peptide can be used in immunogenic
 CC compositions and subunit vaccines for dental caries. These comprise MHC
 CC class II protein-binding GbpB peptides covalently linked with peptide
 CC subunits (preferably from the catalytic domain) of a glucosyltransferase.
 CC The compositions are used in a claimed method of eliciting production of
 CC an antibody in a mammal. Diepitopic or multiepitopic polypeptides can be
 CC prepared synthetically or by recombinant DNA technology. Antibodies
 CC raised against MHC class II binding fragments of GbpB can be used in
 CC passive immunisation.

XX SQ Sequence 20 AA;

Query Match 67.0%; Score 65; DB 7; Length 20;
 Best Local Similarity 100.0%; Pred. No. 0.0013;
 Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 6 TIQGVSAALQTQQA 19
 |||||

Db 6 TIQGVSAALQTQQA 19
 |||||

RESULT 4

ADX37252
 ID ADX37252 standard; peptide; 20 AA.

XX AC ADX37252;

XX DT 21-APR-2005 (first entry)

XX DE Streptococcus mutant glucan binding protein B peptide #9.

KW immunogenicity; immune stimulation; glucan binding protein-B;
 KW microparticle; major histocompatibility complex; tooth disease.

XX Streptococcus mutans.

XX US2005031633-A1.

XX 10-FEB-2005.

XX 09-MAR-2004; 2004US-00797821.

XX 13-APR-1998; 98US-0081550P.

XX 08-JAN-1999; 99US-0115142P.

XX 12-APR-1999; 99US-00290049.

XX 07-MAR-2002; 2002US-0363209P.

XX 08-AUG-2002; 2002US-0402483P.

XX 07-MAR-2003; 2003US-00383930.

XX (SMIT/) SMITH D J.

XX (TAUB/) TAUBMAN M A.

XX Smith DJ, Taubman MA;

XX WPI; 2005-151644/16.

XX Claim 4; SEQ ID NO 9; 73pp; English.

XX The invention relates to a composition comprising a fragment of a glucan binding protein-B (GbpB) that binds to MHC class II protein, and a biocompatible microparticle, useful for producing an antibody (claimed) for immunizing mammals against dental caries.

XX Streptococcus mutans GbpB-derived peptide of the invention.

XX Query Match 67.0%; Score 65; DB 9; Length 20;

XX Best Local Similarity 100.0%; Pred. No. 0.0013;

XX Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 6 TIQGVVSALQTQQA 19

Db 6 TIQGVVSALQTQQA 19

RESULT 5

ADD93649

ID ADD93649 standard; protein; 431 AA.

XX ADD93649;

XX 29-JAN-2004 (first entry)

XX Streptococcus mutans glucan binding protein-B.

XX Glucan binding protein-B; GbpB; vaccine; anticaries; epitope; immunogen.

XX Streptococcus mutans.

XX Key Location/Qualifiers

FT Region 6..25

FT Region /note= "HLA-binding peptide"

FT Region 16..35

FT Region /note= "HLA-binding peptide"

FT Region 33..52

FT Region /note= "HLA-binding peptide"

FT Region 37..56

FT Region /note= "HLA-binding peptide"

FT Region /note= "HLA-binding peptide"

FT Region /note= "HLA-binding peptide"

FT Region /note= "HLA-binding peptide"

FT Region /note= "HLA-binding peptide"

FT Region /note= "HLA-binding peptide"

48..67

/note= "HLA-binding peptide"

52..71

/note= "HLA-binding peptide"

88..107

/note= "HLA-binding peptide"

113..132

/note= "HLA-binding peptide"

117..136

/note= "HLA-binding peptide"

137..156

/note= "HLA-binding peptide"

174..193

/note= "HLA-binding peptide"

194..213

/note= "HLA-binding peptide"

214..233

/note= "HLA-binding peptide"

248..267

/note= "HLA-binding peptide"

289..308

/note= "HLA-binding peptide"

306..325

/note= "HLA-binding peptide"

311..330

/note= "HLA-binding peptide"

349..368

/note= "HLA-binding peptide"

365..384

/note= "HLA-binding peptide"

383..402

/note= "HLA-binding peptide"

403..422

/note= "HLA-binding peptide"

WO2003075845-A2.

18-SEP-2003.

07-MAR-2003; 2003WO-US006962.

07-MAR-2002; 2002US-0363209P.

08-AUG-2002; 2002US-0402483P.

(FORS-) FORSYTH INST.

Smith DJ, Taubman MA;

WPI; 2003-845091/78.

GENBANK; AY046410.

Composition useful as vaccines for dental caries comprises a fragment of a glucan binding protein-B binding to a major histocompatibility complex class II protein.

Claim 6; Page 7; 49pp; English.

The present sequence is the protein sequence of the glucan binding protein-B (GbpB) of Streptococcus mutans strain SK32. The sequence includes a number of human leucocyte antigen (HLA)-binding epitopes. The invention provides immunogenic compositions and vaccines for dental caries. The compositions comprise major histocompatibility complex (MHC) class II protein-binding GbpB peptides, especially HLA-binding peptides, covalently linked with peptide subunits of a glucosyltransferase. The compositions are used in a claimed method of eliciting production of an antibody in a mammal. Dieptopic or multi-epitopic polypeptides can be prepared synthetically or by recombinant DNA technology. Antibodies raised against MHC class II binding fragments of GbpB can be used in passive immunisation.

Sequence 431 AA;

Query Match 67.0%; Score 65; DB 7; Length 431;

Best Local Similarity 100.0%; Pred. No. 0.047; 0; Mismatches 0; Gaps 0;
 Matches 14; Conservative 0; Indels 0; Gaps 0;

QY 6 TIQGVVSALQTQQA 19
 |||||
 Db 57 TIQGVVSALQTQQA 70

RESULT 6
 ADD93650
 ID ADD93650 standard; protein; 431 AA.
 XX
 AC ADD93650;
 XX
 DT 29-JAN-2004 (first entry)
 XX
 DE Streptococcus mutans glucan binding protein-B.
 XX
 KW Glucan binding protein-B; GbpB; vaccine; anticaries; epitope; immunogen.
 XX
 OS Streptococcus mutans.
 XX
 PN WO2003075845-A2.
 XX
 PD 18-SEP-2003.
 XX
 PF 07-MAR-2003; 2003WO-US006962.
 XX
 PR 07-MAR-2002; 2002US-0363209P.
 XX
 PR 08-AUG-2002; 2002US-0402483P.
 XX
 XX (FORS-) FORSYTH INST.
 PA Smith DJ, Taubman MA;
 PI
 XX
 XX WPI; 2003-845091/78.
 XX
 XX GENBANK; AY046411.
 XX
 PT Composition useful as vaccines for dental caries comprises a fragment of
 PT a glucan binding protein-B binding to a major histocompatibility complex
 PT class II protein.
 XX
 PS Claim 5; Page 8; 49pp; English.
 XX
 CC The present sequence is the protein sequence of the glucan binding
 CC protein-B (GbpB) of Streptococcus mutans strain 3VF4. The sequence
 CC includes a number of human leucocyte antigen (HLA)-binding epitopes. The
 CC invention provides immunogenic compositions and vaccines for dental
 CC caries. The compositions comprise major histocompatibility complex (MHC)
 CC class II protein-binding GbpB peptides, especially HLA-binding peptides,
 CC covalently linked with peptide subunits of a glucosyltransferase. The
 CC compositions are used in a claimed method of eliciting production of an
 CC antibody in a mammal. Diepitopic or multiepitopic polypeptides can be
 CC prepared synthetically or by recombinant DNA technology. Antibodies
 CC raised against MHC class II binding fragments of GbpB can be used in
 CC passive immunisation.
 XX
 SQ Sequence 431 AA;

Query Match 67.0%; Score 65; DB 7; Length 431;
 Best Local Similarity 100.0%; Pred. No. 0.047;
 Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 6 TIQGVVSALQTQQA 19
 |||||
 Db 57 TIQGVVSALQTQQA 70

RESULT 7
 ADD93653
 ID ADD93653 standard; protein; 431 AA.
 XX
 AC ADD93653;

XX 29-JAN-2004 (first entry)
 DT
 XX Streptococcus mutans glucan binding protein-B.
 DE
 XX Glucan binding protein-B; GbpB; vaccine; anticaries; epitope; immunogen.
 KW
 XX Streptococcus mutans.
 OS
 XX WO2003075845-A2.
 PN
 XX 18-SEP-2003.
 PD
 XX 07-MAR-2003; 2003WO-US006962.
 PF
 XX 07-MAR-2002; 2002US-0363209P.
 PR
 XX 08-AUG-2002; 2002US-0402483P.
 PR
 XX (FORS-) FORSYTH INST.
 PA
 XX Smith DJ, Taubman MA;
 PI
 XX
 XX WPI; 2003-845091/78.
 DR
 XX GENBANK; AY046411.
 DR
 XX Composition useful as vaccines for dental caries comprises a fragment of
 PT a glucan binding protein-B binding to a major histocompatibility complex
 PT class II protein.
 PT
 XX
 PS Claim 5; Page 8-9; 49pp; English.
 PS
 XX
 CC The present sequence is the protein sequence of the glucan binding
 CC protein-B (GbpB) of Streptococcus mutans strain 5M3. The sequence
 CC includes a number of human leucocyte antigen (HLA)-binding epitopes. The
 CC invention provides immunogenic compositions and vaccines for dental
 CC caries. The compositions comprise major histocompatibility complex (MHC)
 CC class II protein-binding GbpB peptides, especially HLA-binding peptides,
 CC covalently linked with peptide subunits of a glucosyltransferase. The
 CC compositions are used in a claimed method of eliciting production of an
 CC antibody in a mammal. Diepitopic or multiepitopic polypeptides can be
 CC prepared synthetically or by recombinant DNA technology. Antibodies
 CC raised against MHC class II binding fragments of GbpB can be used in
 CC passive immunisation.
 CC
 XX
 SQ Sequence 431 AA;

Query Match 67.0%; Score 65; DB 7; Length 431;
 Best Local Similarity 100.0%; Pred. No. 0.047;
 Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 6 TIQGVVSALQTQQA 19
 |||||
 Db 57 TIQGVVSALQTQQA 70

RESULT 8
 ADX37272
 ID ADX37272 standard; protein; 431 AA.
 XX
 AC ADX37272;
 XX
 DT 21-APR-2005 (first entry)
 DT
 XX Streptococcus mutant glucan binding protein B variant #1.
 DE
 XX immunogenicity; immune stimulation; glucan binding protein-B;
 KW microparticle; major histocompatibility complex; tooth disease.
 KW
 XX Streptococcus mutans.
 OS
 XX US2005031633-A1.
 PN
 XX 10-FEB-2005.
 PD

CC fragment binds to a major histocompatibility complex (MHC) class II
CC protein. The composition is useful for producing an antibody for
CC immunizing mammals against dental caries. This sequence corresponds to a
CC Streptococcus mutans GbpB protein of the invention.

XX Sequence 431 AA;

Query Match 67.0%; Score 65; DB 9; Length 431;
Best Local Similarity 100.0%; Pred. No. 0.047;
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 6 TIQQGVSAALQTQQA 19
|||||
DB 57 TIQQGVSAALQTQQA 70

RESULT 11

AEB91500
ID AEB91500 standard; protein; 431 AA.

AC AEB91500;

XX 20-OCT-2005 (first entry)

XX Microbial pathogen adhesin protein sequence, SEQ ID NO:210.

XX algorithm; adhesin; pharmaceutical; vaccine; drug screening;
KW bordetella pertussis infection; antibacterial; pneumonia;
KW antiinflammatory; respiratory-gen.; gastric ulcer; antiulcer;
KW gastrointestinal-gen.; urinary tract infection; antimicrobial; uropathic.

XX Streptococcus mutans.

XX WO2005076010-A2.

XX 18-AUG-2005.

XX 07-FEB-2005; 2005WO-IN000037.

XX 06-FEB-2004; 2004IN-DE000173.

XX 20-JUL-2004; 2004US-0589227P.

XX (COUL) COUNCIL SCI & IND RES SOUTH AFRICA.

XX Sachdeva G, Kumar K, Jain P, Brahmachari SK, Ramachandran S;

XX WPI; 2005-597835/61.

XX Computational method for identifying adhesin and adhesin like molecules,
PT comprises computing sequence-based attributes of protein sequences using
PT neural network software and training an artificial neural network.

XX Claim 16; SEQ ID NO 210; 402pp; English.

XX The present invention relates to a computational method (M1) for
CC identifying adhesin and adhesin-like proteins, by computing the sequence-
CC based attributes of protein sequences using five attribute modules of a
CC neural network software, training an artificial neural network (ANN) for
CC each of the computed five attributes, and identifying the adhesin and
CC adhesin-like proteins having probability of being an adhesin (Pad) as
CC equal or greater than 0.51. Also claimed is a set of 274 annotated genes
CC encoding adhesin and adhesin-like proteins, having 274 fully defined 162-
CC 1151 base pair (SEQ ID NO: 385-658) sequences; a set of 105 hypothetical
CC genes encoding adhesin and adhesin-like proteins, having 105 fully
CC defined 306-15876 base pairs (SEQ ID NO: 659-763); a set of 279 annotated
CC adhesin and adhesin-like proteins, having 279 fully defined 53-3716 base
CC pair (SEQ ID NO: 1-279); a set of 105 hypothetical adhesin and adhesin-
CC like proteins, having 105 fully defined 106-5291 base pair (SEQ ID NO:
CC 280-384) sequences; and a fully connected multilayer feed forward ANN (I)
CC based on (M1). (M1) is useful for identifying adhesin and adhesin-like
CC proteins, of therapeutic potential, and identifying and short-listing
CC proteins for further testing in development of new vaccine formulations
CC to eliminate diseases caused by various pathogenic organisms. (M1) is

CC useful for identifying putative adhesins that are important in drug
CC discovery and preventing therapeutics for whooping cough, pneumonia,
CC gastric ulcer and urinary tract infections. (M1) identifies adhesins from
CC distantly related organisms, and from bacteria belonging to a wide
CC phylogenetic spectrum. (M1) is capable of predicting adhesive nature of
CC unique proteins. The present sequence is a microbial pathogen adhesin
CC protein sequence.

XX Sequence 431 AA;

Query Match 67.0%; Score 65; DB 9; Length 431;
Best Local Similarity 100.0%; Pred. No. 0.047;
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 6 TIQQGVSAALQTQQA 19
|||||
DB 57 TIQQGVSAALQTQQA 70

RESULT 12

ADD93651
ID ADD93651 standard; protein; 432 AA.

XX AC ADD93651;

XX 29-JAN-2004 (first entry)

XX Streptococcus mutans glucan binding protein-B.

XX Glucan binding protein-B; GbpB; vaccine; anticaries; epitope; immunogen.

XX Streptococcus mutans.

XX WO2003075845-A2.

XX 18-SEP-2003.

XX 07-MAR-2003; 2003WO-US006962.

XX 07-MAR-2002; 2002US-0363209P.

XX 08-AUG-2002; 2002US-0402483P.

XX (FORS-) FORSYTH INST.

XX Smith DJ, Taubman MA;

XX WPI; 2003-845091/78.

XX GENBANK; AY046412.

XX Composition useful as vaccines for dental caries comprises a fragment of
PT a glucan binding protein-B binding to a major histocompatibility complex
PT class II protein.

XX Claim 5; Page 8; 49pp; English.

XX The present sequence is the protein sequence of the glucan binding
CC protein-B (GbpB) of Streptococcus mutans strain 15P2. The sequence
CC includes a number of human leucocyte antigen (HLA) -binding epitopes. The
CC invention provides immunogenic compositions and vaccines for dental
CC caries. The compositions comprise major histocompatibility complex (MHC)
CC class II protein-binding GbpB peptides, especially HLA-binding peptides,
CC covalently linked with peptide subunits of a glucosyltransferase. The
CC compositions are used in a claimed method of eliciting production of an
CC antibody in a mammal. Dieptopic or multiepitopic polypeptides can be
CC prepared synthetically or by recombinant DNA technology. Antibodies
CC raised against MHC class II binding fragments of GbpB can be used in
CC passive immunisation.

XX Sequence 432 AA;

Query Match 67.0%; Score 65; DB 7; Length 432;
Best Local Similarity 100.0%; Pred. No. 0.047;
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 6 TTIQGVSAALQTQQA 19
 Db 57 TTIQGVSAALQTQQA 70

RESULT 13
 ADD93652
 ID ADD93652 standard; protein; 432 AA.
 XX AC ADD93652;
 XX DT 29-JAN-2004 (first entry)
 XX DE Streptococcus mutans glucan binding protein-B.
 XX KW Glucan binding protein-B; GbpB; vaccine; anticaries; epitope; immunogen.
 XX OS Streptococcus mutans.
 XX PN WO2003075845-A2.
 XX PD 18-SEP-2003.
 XX PF 07-MAR-2003; 2003WO-US006962.
 XX PR 07-MAR-2002; 2002US-0363209P.
 XX PR 08-AUG-2002; 2002US-0402483P.
 XX PA (FORSYTH INST.
 XX PI Smith DJ, Taubman MA;
 XX PS WPI; 2003-845091/78.
 XX DR GENBANK; AY046413.
 XX DT Composition useful as vaccines for dental caries comprises a fragment of a glucan binding protein-B binding to a major histocompatibility complex class II protein.
 XX PS Claim 5; Page 8; 49pp; English.
 XX CC The present sequence is the protein sequence of the glucan binding protein-B (GbpB) of Streptococcus mutans strain 3SM1. The sequence includes a number of human leucocyte antigen (HLA)-binding epitopes. The invention provides immunogenic compositions and vaccines for dental caries. The compositions comprise major histocompatibility complex (MHC) class II protein-binding GbpB peptides, especially HLA-binding peptides, covalently linked with peptide subunits of a glucosyltransferase. The compositions are used in a claimed method of eliciting production of an antibody in a mammal. Dieptopic or multiepitopic polypeptides can be prepared synthetically or by recombinant DNA technology. Antibodies raised against MHC class II binding fragments of GbpB can be used in passive immunisation.

QY 6 TTIQGVSAALQTQQA 19
 Db 57 TTIQGVSAALQTQQA 70

RESULT 14
 ADX37274
 ID ADX37274 standard; protein; 432 AA.
 XX AC ADX37274;
 XX DT 21-APR-2005 (first entry)

QY 6 TTIQGVSAALQTQQA 19
 Db 57 TTIQGVSAALQTQQA 70

RESULT 15
 ADX37275
 ID ADX37275 standard; protein; 432 AA.
 XX AC ADX37275;
 XX DT 21-APR-2005 (first entry)
 XX DE Streptococcus mutans glucan binding protein B variant #4.
 XX KW immunogenicity; immune stimulation; glucan binding protein-B; microparticle; major histocompatibility complex; tooth disease.
 XX OS Streptococcus mutans.
 XX PN US2005031633-A1.
 XX PD 10-FEB-2005.
 XX PF 09-MAR-2004; 2004US-00797821.

XX Streptococcus mutant glucan binding protein B variant #3.
 XX immunogenicity; immune stimulation; glucan binding protein-B;
 KW microparticle; major histocompatibility complex; tooth disease.
 XX OS Streptococcus mutans.
 XX PN US2005031633-A1.
 XX PD 10-FEB-2005.
 XX PF 09-MAR-2004; 2004US-00797821.
 XX PR 13-APR-1998; 98US-0081550P.
 XX PR 08-JAN-1999; 99US-0115142P.
 XX PR 12-APR-1999; 99US-0029004P.
 XX PR 07-MAR-2002; 2002US-0363209P.
 XX PR 08-AUG-2002; 2002US-0402483P.
 XX PR 07-MAR-2003; 2003US-00383930.
 XX PA (SMITH) SMITH D J.
 XX PA (TAUB/) TAUBMAN M A.
 XX PI Smith DJ, Taubman MA;
 XX DR WPI; 2005-151644/16.
 XX DT New composition comprising a fragment of a glucan binding protein-B (GbpB) that binds to MHC class II protein, and a biocompatible microparticle, useful for producing an antibody (claimed) for immunizing mammals against dental caries.
 XX PS Claim 3; SEQ ID NO 31; 73pp; English.
 XX CC The invention relates to a composition comprising a fragment of a glucan binding protein-B (GbpB) and a biocompatible microparticle, where the fragment binds to a major histocompatibility complex (MHC) class II protein. The composition is useful for producing an antibody for immunizing mammals against dental caries. This sequence corresponds to a Streptococcus mutans GbpB protein of the invention.

QY 6 TTIQGVSAALQTQQA 19
 Db 57 TTIQGVSAALQTQQA 70

Query Match 67.0%; Score 65; DB 9; Length 432;
 Best Local Similarity 100.0%; Pred. No. 0.047;
 Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 6 TTIQGVSAALQTQQA 19
 Db 57 TTIQGVSAALQTQQA 70

RESULT 15
 ADX37275
 ID ADX37275 standard; protein; 432 AA.
 XX AC ADX37275;
 XX DT 21-APR-2005 (first entry)
 XX DE Streptococcus mutant glucan binding protein B variant #4.
 XX KW immunogenicity; immune stimulation; glucan binding protein-B; microparticle; major histocompatibility complex; tooth disease.
 XX OS Streptococcus mutans.
 XX PN US2005031633-A1.
 XX PD 10-FEB-2005.
 XX PF 09-MAR-2004; 2004US-00797821.

```

XX      13-APR-1998;      98US-0081550P.
PR      08-JAN-1999;      99US-0115142P.
PR      12-APR-1999;      99US-00290049.
PR      07-MAR-2002;      2002US-0363209P.
PR      08-AUG-2002;      2002US-0402483P.
PR      07-MAR-2003;      2003US-00383930.
XX      (SMIT/) SMITH D J.
XX      (TAUB/) TAUBMAN M A.
XX
XX      Smith DJ, Taubman MA;
XX
XX      WPI; 2005-151644/16.
XX
XX      New composition comprising a fragment of a glucan binding protein-B
PT      (GbpB) that binds to MHC class II protein, and a biocompatible
PT      microparticle, useful for producing an antibody (claimed) for immunizing
PT      mammals against dental caries.
XX
XX      Claim 3; SEQ ID NO 32; 73pp; English.
XX
XX      The invention relates to a composition comprising a fragment of a glucan
CC      binding protein-B (GbpB) and a biocompatible microparticle, where the
CC      fragment binds to a major histocompatibility complex (MHC) class II
CC      protein. The composition is useful for producing an antibody for
CC      immunizing mammals against dental caries. This sequence corresponds to a
CC      Streptococcus mutans GbpB protein of the invention.
XX
XX      Sequence 432 AA;
SQ
      Query Match      67.0%; Score 65; DB 9; Length 432;
      Best Local Similarity 100.0%; Pred. No. 0.047;
      Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
      QY      6 TIOGQVSALQTQQA 19
      DB      57 TIOGQVSALQTQQA 70
Search completed: January 25, 2006, 19:01:40
Job time : 69.2 secs

```


C:Genetics:
A:Gene: Cj0859c
C:Superfamily: Campylobacter jejuni hypothetical protein Cj0859c

Query Match 46.4%; Score 45; DB 2; Length 142;
Best Local Similarity 64.3%; Pred. No. 4.2;
Matches 9; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 6 TTIQGVVSALQTQQA 19
||| : ||| : ||| :
Db 129 TTIQAIQELQSQQA 142

RESULT 3
AB0551
exonuclease SbcC [imported] - Salmonella enterica subsp. enterica serovar Typhi (strain
C:Species: Salmonella enterica subsp. enterica serovar Typhi
A:Title: This species has also been called Salmonella typhi
C:Date: 09-Nov-2001 #sequence_revision 09-Nov-2001 #text_change 18-Nov-2002
C:Accession: AB0551
R:Parkhill, J.; Dougan, G.; James, K.D.; Thomson, N.R.; Pickard, D.; Wain, J.; Churcher,
th, T.; Conner, P.; Cronin, A.; Davis, P.; Davies, R.M.; Dowd, L.; White, N.; Farrar,
S.; Moule, S.; O'Gaora, P.
Nature 413, 848-852, 2001
A:Authors: Parry, C.; Quail, M.; Rutherford, K.; Simmonds, M.; Skelton, J.; Stevens, K.;
A:Title: Complete genome sequence of a multiple drug resistant Salmonella enterica serov
A:Reference number: AB0502; MUID:21534947; PMID:11677608
A:Accession: AB0551
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-1034 <PAR>
A:Cross-references: UNIPARC:UPI00000CDAPF; GB:AL513382; PIDN:CAD08850.1; PID:G16501663;
C:Genetics:
A:Gene: STY0429
C:Superfamily: sbcC protein

Query Match 46.4%; Score 45; DB 2; Length 1034;
Best Local Similarity 42.1%; Pred. No. 34;
Matches 8; Conservative 5; Mismatches 6; Indels 0; Gaps 0;

QY 1 KHKLIITIQGVVSALQTQQA 19
||| : ||| : ||| :
Db 425 RHRLAALQGIPLPKQRQA 443

RESULT 4
D69617
DNA polymerase III (alpha subunit) dnaE - Bacillus subtilis
C:Species: Bacillus subtilis
C:Date: 05-Dec-1997 #sequence_revision 05-Dec-1997 #text_change 09-Jul-2004
C:Accession: D69617
R:Kunst, F.; Ogasawara, N.; Moszer, I.; Albertini, A.M.; Alloni, G.; Azevedo, V.; Berte
C.; Bron, S.; Brouillet, S.; Brusch, C.V.; Caldwell, B.; Capuano, V.; Carter, N.M.; Ch
A.; Ehrlich, S.D.; Emerson, P.T.; Entian, K.D.; Errington, J.; Fabbre, C.; Ferrari, E.
Nature 390, 249-256, 1997
A:Authors: Foulger, D.; Fritz, C.; Fujita, M.; Fujita, Y.; Fuma, S.; Galizzi, A.; Gall
iech, J.; Harwood, C.R.; Henaut, A.; Hilbert, H.; Holsappel, S.; Hosono, S.; Hullo, M.P.
Koetter, P.; Koningsstein, G.; Krogh, S.; Kumano, M.; Kurita, K.; Lapidus, A.; Lardinois,
A:Authors: Lauber, J.; Lazarevic, V.; Lee, S.M.; Levine, A.; Liu, H.; Maguda, S.; Maue
Y, M.; Ogawa, K.; Ogiwara, A.; Oudega, B.; Park, S.H.; Parro, V.; Pohl, T.M.; Portetel
Rieger, M.; Rivolta, C.; Rocha, E.; Roche, B.; Rose, M.; Sadaie, Y.; Sato, T.; Scanlon,
A:Authors: Schleich, S.; Schroeter, R.; Scoffone, F.; Sekiguchi, J.; Sekowska, A.; Seron
akeuchi, M.; Tamakoshi, A.; Tanaka, T.; Terpstra, P.; Tognoni, A.; Tosato, V.; Uchiyama,
T.; Winters, P.; Wipat, A.; Yamamoto, H.; Yamane, K.; Yasumoto, K.; Yata, K.; Yoshida, K
A:Authors: Yoshikawa, H.F.; Zumstein, E.; Yoshikawa, H.; Danchin, A.
A:Title: The complete genome sequence of the Gram-positive bacterium Bacillus subtilis.
A:Reference number: A69580; MUID:98044033; PMID:9384377
A:Accession: D69617
A:Status: preliminary; nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-1115 <KUN>
A:Cross-references: UNIPROT:O34623; UNIPARC:UPI00000608C4; GB:Z99118; GB:AL009126; NID:9
A:Experimental source: strain 168

C:Genetics:
A:Gene: dnaE
C:Superfamily: DNA-directed DNA polymerase III alpha chain

Query Match 46.4%; Score 45; DB 2; Length 1115;
Best Local Similarity 58.8%; Pred. No. 36;
Matches 10; Conservative 2; Mismatches 5; Indels 0; Gaps 0;

QY 3 KLITIQGVVSALQTQQA 19
||| : ||| : ||| :
Db 933 KQLTAQGAIVSLQAQA 949

RESULT 5
D70673
hypothetical protein Rv2983 - Mycobacterium tuberculosis (strain H37RV)
C:Species: Mycobacterium tuberculosis
C:Date: 17-Jul-1998 #sequence_revision 17-Jul-1998 #text_change 09-Jul-2004
C:Accession: D70673
R:Cole, S.T.; Brosch, R.; Parkhill, J.; Garnier, T.; Churcher, C.; Harris, D.; Gordon, S.
; Connor, R.; Davies, R.; Devlin, K.; Feldwell, T.; Gentles, S.; Hamlin, N.; Holroyd, S.
Rajandream, M.A.; Rogers, J.; Rutter, S.; Seeger, K.; Skelton, S.; Squares, S.
Nature 393, 537-544, 1998
A:Authors: Sgares, R.; Sulston, J.E.; Taylor, K.; Whitehead, S.; Barrell, B.G.
A:Title: Deciphering the biology of Mycobacterium tuberculosis from the complete genome
A:Reference number: A70500; MUID:98295987; PMID:9634230
A:Accession: D70673
A:Status: preliminary; nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-214 <COL>
A:Cross-references: UNIPROT:P95112; UNIPARC:UPI00000D0F78; GB:Z83018; GB:AL123456; NID:
A:Experimental source: strain H37RV
C:Genetics:
A:Gene: Rv2983

Query Match 44.3%; Score 43; DB 2; Length 214;
Best Local Similarity 46.7%; Pred. No. 14;
Matches 7; Conservative 5; Mismatches 3; Indels 0; Gaps 0;

QY 4 LITIQGVVSALQTQQ 18
||| : ||| : ||| :
Db 110 IVVLQGLPALQTQE 124

RESULT 6
T45430
hypothetical protein [imported] - Mycobacterium leprae
C:Species: Mycobacterium leprae
C:Date: 31-Jan-2000 #sequence_revision 31-Jan-2000 #text_change 09-Jul-2004
C:Accession: T45430
R:Parkhill, J.; Barrell, B.G.; Rajandream, M.A.
submitted to the EMBL Data Library, September 1997
A:Reference number: Z16918
A:Accession: T45430
A:Status: preliminary; translated from GB/EMBL/DDBJ
A:Molecule type: DNA
A:Residues: 1-216 <PAR>
A:Cross-references: UNIPROT:O33128; UNIPARC:UPI00000C60E3; EMBL:Z99263; PIDN:CABL6452.1
A:Experimental source: cosmid B637
C:Genetics:
A:Note: MLCB637.37C

Query Match 44.3%; Score 43; DB 2; Length 216;
Best Local Similarity 46.7%; Pred. No. 14;
Matches 7; Conservative 5; Mismatches 3; Indels 0; Gaps 0;

QY 4 LITIQGVVSALQTQQ 18
||| : ||| : ||| :
Db 112 IVVLQGLPALQTQE 126

RESULT 7
D95309

C;Species: Arabidopsis thaliana (mouse-ear cress)
C;Date: 12-Feb-1999 #sequence_revision 12-Feb-1999 #text_change 09-Jul-2004
C;Accession: T00931; C84851
R;Rounsley, S.D.; Lin, X.; Ketchum, K.A.; Crosby, M.L.; Brandon, R.C.; Sykes, S.M.; Kaul
submitted to the EMBL Data Library, December 1997
A;Description: Arabidopsis thaliana chromosome II BAC T24P15 genomic sequence.
A;Reference number: Z14212
A;Accession: T00931
A;Status: translated from GB/EMBL/DBJ
A;Molecule type: DNA
A;Residues: 1-234 <ROU>
A;Cross-references: UNIPROT:O48529; UNIPARC:UPI00000A022B; EMBL:AC002561; NID:g2673901;
A;Experimental source: cultivar Columbia
R;Lin, X.; Kaul, S.; Rounsley, S.D.; Shea, T.P.; Benito, M.I.; Town, C.D.; Fujii, C.Y.;
M.; Koo, H.; Moffat, K.S.; Cronin, L.A.; Shen, M.; VanAken, S.E.; Umavam, L.; Tallon, L.;
euss, D.; Nierman, W.C.; White, O.; Eisen, J.A.; Salzberg, S.L.; Fraser, C.M.; Venter, J.
Nature 402, 761-768, 1999
A;Title: Sequence and analysis of chromosome 2 of the plant Arabidopsis thaliana.
A;Reference number: A84420; MUID:20083487; PMID:10617197
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-234 <STO>
A;Cross-references: UNIPARC:UPI00000A022B; GB:AE002093; NID:g2673913; PIDN:AA886647.1; Q
C;Genetics:
A;Gene: At3g42220; T24P15.13
A;Map position: 2
A;Introns: 16/2; 100/1; 150/2; 179/1; 214/1

—

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2006 Compugen Ltd.

OM protein - protein search, using sw model

Run on: January 25, 2006, 18:48:24 ; Search time 70.3 Seconds
(without alignments)
200.719 Million cell updates/sec

Title: US-10-797-821-2
Perfect score: 97
Sequence: 1 KHKLITIQGVSAALQTQQA 20

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 2166443 seqs, 705528306 residues

Total number of hits satisfying chosen parameters: 2166443

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Uniprot 05.80:*
1: uniprot_sprot:*
2: uniprot_trembl:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	91	93.8	211	2 Q9ZAS7_STRMU	Q9zas7 streptococc
2	65	67.0	431	2 Q938V0_STRMU	Q938v0 streptococc
3	65	67.0	431	2 Q938V3_STRMU	Q938v3 streptococc
4	65	67.0	431	2 Q9AG98_STRMU	Q9ag98 streptococc
5	65	67.0	431	2 Q8DMW3_STRMU	Q8dmw3 streptococc
6	65	67.0	432	2 Q938V1_STRMU	Q938v1 streptococc
7	65	67.0	432	2 Q938V2_STRMU	Q938v2 streptococc
8	51	52.6	130	2 Q5KRH3_CORGL	Q5kth3 corynebacte
9	48	49.5	448	2 Q6ALV1_DESPS	Q6alv1 desulfotale
10	47	48.5	529	2 Q6BUA6_DEBHA	Q6bua6 debaryomyce
11	46	47.4	223	2 Q4T8Q5_TETNG	Q4t8q5 tetraodon n
12	46	47.4	465	2 Q82LB1_STRAW	Q82lb1 streptomyce
13	46	47.4	592	2 Q5UTU0_CANAL	Q5ut0 candida alb
14	46	47.4	636	2 Q9LP22_ARATH	Q9lp22 arabidopsis
15	46	47.4	794	2 Q4SK16_TETNG	Q4sk16 tetraodon n
16	45	46.4	142	2 Q5HUT7_CAMJR	Q5hut7 campylobact
17	45	46.4	142	2 Q9PB64_CAMJE	Q9pb64 campylobact
18	45	46.4	408	2 Q6SUD8_MANSM	Q6sud8 mannelmia
19	45	46.4	447	2 Q9AKA4_STRAG	Q9aka4 streptococc
20	45	46.4	447	2 Q8E2H1_STRAS	Q8e2h1 streptococc
21	45	46.4	447	2 Q8E7X9_STRAS	Q8e7x9 streptococc
22	45	46.4	458	2 Q7XBT7_ORYSA	Q7xbt7 oryza lativ
23	45	46.4	458	2 Q9FQO2_ORYSA	Q9fqo2 oryza lativ
24	45	46.4	1034	2 Q8ZBY6_SALTI	Q8zby6 salmonella
25	45	46.4	1036	2 Q83S29_SALTI	Q83sz9 salmonella
26	45	46.4	1046	2 Q57SG9_SALCH	Q57sg9 salmonella
27	45	46.4	1046	2 Q8ZRE3_SALTY	Q8zre3 salmonella
28	45	46.4	1046	2 Q5PFU4_SALPA	Q5pfu4 salmonella
29	45	46.4	1115	1 DP03A_BACSU	Q34623 bacillus su
30	45	46.4	1489	2 Q54X15_DICDI	Q54x15 dictyosteli
31	44.5	45.9	535	2 Q8KWC2_RUEHOB	Q8kwc2 ruegeria sp

RESULT 1
Q9ZAS7_STRMU
ID Q9ZAS7_STRMU PRELIMINARY; PRT; 211 AA.
AC Q9ZAS7;
DT 01-MAY-1999 (TrEMBLrel. 10, Created)
DT 01-MAY-1999 (TrEMBLrel. 10, Last sequence update)
DT 01-OCT-2002 (TrEMBLrel. 22, Last annotation update)
DE Putative secreted protein (Fragment).
OS Streptococcus mutans.
OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
OC Streptococcus.
OX NCBI_TaxID=1309;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=GS-5 Kuramitsu;
RA Peruzzi F., Piggot P.J., Daneo-Moore L.;
RL Submitted (NOV-1996) to the EMBL/GenBank/DBJ databases.
DR EMBL; U78607; AAD00288.1; -; Genomic_DNA.
FT NON_TER 211
SQ SEQUENCE 211 AA; 22803 MW; 4ACE331159CFAFC6 CRC64;

Query Match 93.8%; Score 91; DB 2; Length 211;
Best Local Similarity 100.0%; Pred. No. 9.9e-07;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KHKLITIQGVSAALQTQQA 19
|||||
DB 52 KHKLITIQGVSAALQTQQA 70
|||||

RESULT 2
Q938V0_STRMU
ID Q938V0_STRMU PRELIMINARY; PRT; 431 AA.
AC Q938V0;
DT 01-DEC-2001 (TrEMBLrel. 19, Created)
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Glucan-binding protein B.
OS Streptococcus mutans.
OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
OC Streptococcus.
OX NCBI_TaxID=1309;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=5SM3;
RA Jin S., Duncan M.J., Taubman M.A., Smith D.J.;
RT "Cloning of the gppB gene from Streptococcus mutans.";
RL J. Dent. Res. 79:224-224 (2000).
[2]
RC STRAIN=5SM3;
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=5SM3;
RX MEDLINE=21481971; PubMed=11598068;
DOI=10.1128/JAI.69.11.6931-6941.2001;
RA Mattos-Graner R.O., Jin S., King W.F., Chen T., Smith D.J.,

32 44 45.4 291 2 Q838T4_ENTFA Q838t4 enterococcu
33 44 45.4 410 2 Q9F4K8_RHILU Q9f4k8 rhizobium 1
34 44 45.4 410 2 Q6FZS2_BARQU Q6fzs2 bartonella
35 44 45.4 412 1 PEPT_PASMU Q9cp05 pasteurilla
36 44 45.4 457 2 Q8FBE3_ECOL6 Q8fbe3 escherichia
37 44 45.4 474 2 Q5MGK4_STRT2 Q5mgk4 streptococc
38 44 45.4 485 2 Q5M212_STRT1 Q5m212 streptococc
39 44 45.4 487 2 Q64D30_9ARCH Q64d30 uncultured
40 44 45.4 780 2 Q06271_XENLA Q06271 xenopus lae
41 44 45.4 788 1 LRNF2_MOUSE LRNF2 mouse
42 44 45.4 789 1 LRNF2_HUMAN Q9ulh4 homo sapien
43 44 45.4 789 1 LRNF2_HUMAN Q9be71 macaca fasc
44 44 45.4 1184 2 Q7VEE2_PROMA Q7vee2 prochloroco
45 44 45.4 1319 2 Q5B761_EMENI Q5b761 aspergillus

```

RA Duncan M.J.;
RT "Cloning of the Streptococcus mutans gene encoding glucan binding
RT protein B and analysis of genetic diversity and protein production in
RT clinical isolates.";
RL Infect. Immun. 69:6931-6941(2001).
DR EMBL; AY046414; AAK94504.1; -; Genomic_DNA.
DR InterPro; IPR007921; CHAP.
DR InterPro; IPR009148; Siba.
DR Pfam; PF05257; CHAP; 1.
DR PRINTS; PR01852; SIBAPROTEIN.
DR PROSITE; PS50911; CHAP; 1.
SQ SEQUENCE 431 AA; 44650 MW; 05D38D8B8C4609F CRC64;

Query Match 67.0%; Score 65; DB 2; Length 431;
Best Local Similarity 100.0%; Pred. No. 0.046;
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 6 TIOGQVSALQTQQA 19
DB 57 TIOGQVSALQTQQA 70

RESULT 3
Q938V3_STRMU PRELIMINARY; PRT; 431 AA.
AC Q938V3;
DT 01-DEC-2001 (TrEMBLrel. 19, Created)
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Glucan-binding protein B.
OS Streptococcus mutans.
OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
OC Streptococcus.
OX NCBI_TaxID=1309;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=SU32;
RA Jin S., Duncan M.J., Taubman M.A., Smith D.J.;
RT "Cloning of the gbpB gene from Streptococcus mutans.";
RL J. Dent. Res. 79:224-224(2000).
RN [2]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=SU32;
RX MEDLINE=21481971; PubMed=11598068;
RX DOI=10.1128/IAI.69.11.6931-6941.2001;
RA Mattos-Graner R.O., Jin S., King W.F., Chen T., Smith D.J.,
RA Duncan M.J.;
RT "Cloning of the Streptococcus mutans gene encoding glucan binding
RT protein B and analysis of genetic diversity and protein production in
RT clinical isolates.";
RL Infect. Immun. 69:6931-6941(2001).
DR EMBL; AY046410; AAK94500.1; -; Genomic_DNA.
DR InterPro; IPR007921; CHAP.
DR InterPro; IPR009148; Siba.
DR Pfam; PF05257; CHAP; 1.
DR PRINTS; PR01852; SIBAPROTEIN.
DR PROSITE; PS50911; CHAP; 1.
SQ SEQUENCE 431 AA; 44620 MW; 464FE3B563PB7E51 CRC64;

Query Match 67.0%; Score 65; DB 2; Length 431;
Best Local Similarity 100.0%; Pred. No. 0.046;
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 6 TIOGQVSALQTQQA 19
DB 57 TIOGQVSALQTQQA 70

RESULT 4
Q9AG98_STRMU PRELIMINARY; PRT; 431 AA.
ID Q9AG98_STRMU PRELIMINARY; PRT; 431 AA.
AC Q9AG98;
DT 01-JUN-2001 (TrEMBLrel. 17, Created)

```

```

DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
DT 01-FEB-2005 (TrEMBLrel. 29, Last annotation update)
DE Immunodominant glycoprotein IBG-60 (Glucan-binding protein B).
GN Name=saga;
OS Streptococcus mutans.
OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
OC Streptococcus.
OX NCBI_TaxID=1309;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=GS-5;
RX MEDLINE=21153617; PubMed=11254612;
RX DOI=10.1128/IAI.69.4.2493-2501.2001;
RA Chia J.S., Lee Y.Y., Huang P.T., Chen J.Y.;
RT "Identification of stress-responsive genes in Streptococcus mutans by
RT differential display reverse transcription-PCR.";
RL Infect. Immun. 69:2493-2501(2001).
RN [2]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=GS-5;
RX MEDLINE=21481977; PubMed=11598074;
RX DOI=10.1128/IAI.69.11.6987-6998.2001;
RA Chia J.S., Chang L.Y., Shun C.T., Chang Y.Y., Chen J.Y.;
RT "A 60-kilodalton immunodominant glycoprotein is essential for cell
RT wall integrity and the maintenance of cell shape in Streptococcus
RT mutans.";
RL Infect. Immun. 69:6987-6998(2001).
RN [3]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=GS-5;
RA Chia J.S., Chang L.Y., Lee Y.Y., Chen J.-Y.;
RL Submitted (NOV-2001) to the EMBL/GenBank/DBJ databases.
RN [4]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=3VF4;
RA Jin S., Duncan M.J., Taubman M.A., Smith D.J.;
RT "Cloning of the gbpB gene from Streptococcus mutans.";
RL J. Dent. Res. 79:224-224(2000).
RN [5]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=3VF4;
RX MEDLINE=21481971; PubMed=11598068;
RX DOI=10.1128/IAI.69.11.6931-6941.2001;
RA Mattos-Graner R.O., Jin S., King W.F., Chen T., Smith D.J.,
RA Duncan M.J.;
RT "Cloning of the Streptococcus mutans gene encoding glucan binding
RT protein B and analysis of genetic diversity and protein production in
RT clinical isolates.";
RL Infect. Immun. 69:6931-6941(2001).
DR EMBL; AF338445; AAK08104.1; -; Genomic_DNA.
DR EMBL; AY046411; AAK94501.1; -; Genomic_DNA.
DR InterPro; IPR007921; CHAP.
DR InterPro; IPR009148; Siba.
DR Pfam; PF05257; CHAP; 1.
DR PRINTS; PR01852; SIBAPROTEIN.
DR PROSITE; PS50911; CHAP; 1.
SQ SEQUENCE 431 AA; 44592 MW; 3EBE21FC5E47232E CRC64;

Query Match 67.0%; Score 65; DB 2; Length 431;
Best Local Similarity 100.0%; Pred. No. 0.046;
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 6 TIOGQVSALQTQQA 19
DB 57 TIOGQVSALQTQQA 70

RESULT 5
Q8DWM3_STRMU PRELIMINARY; PRT; 431 AA.
ID Q8DWM3_STRMU PRELIMINARY; PRT; 431 AA.
AC Q8DWM3;
DT 01-MAR-2003 (TrEMBLrel. 23, Created)
DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)

```


DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
 DE Putative secreted antigen GbpB/SagA, putative peptidoglycan
 DE hydrolase.
 GN Name=gbpB; OrderedLocusNames=SMU.22;
 OS Streptococcus mutans.
 OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
 OC Streptococcus.
 OX NCBI_TaxID=1309;
 RN [1]
 RP NUCLEOTIDE SEQUENCE.
 RC STRAIN=UAI59 / ATCC 700610 / Serotype c;
 RX MEDLINE=22295063; PubMed=12397186; DOI=10.1073/pnas.172501299;
 RA Ajdic D.J., McShan W.M., McLaughlin R.E., Savic G., Chang J.,
 RA Carson M.B., Primeaux C., Tian R., Kenton S., Jia H.G., Lin S.P.,
 RA Qian Y., Li S., Zhu H., Najjar F.Z., Lai H., White J., Roe B.A.,
 RA Ferretti J.J.;
 RT "Genome sequence of Streptococcus mutans UAI59, a cariogenic dental
 RT pathogen.";
 RL Proc. Natl. Acad. Sci. U.S.A. 99:14434-14439(2002).
 DR EMBL; AR014855; AN57811.1; -; Genomic_DNA.
 DR InterPro; IPR007921; CHAP.
 DR InterPro; IPR009148; Siba.
 DR Pfam; PF05257; CHAP; 1.
 DR PRINTS; PR01852; SIBAPROTEIN.
 DR PROSITE; PS50911; CHAP; 1.
 KW Complete proteome.
 SQ SEQUENCE 431 AA; 44620 MW; 2D1CA685248CCD3B CRC64;

Query Match 67.0%; Score 65; DB 2; Length 431;
 Best Local Similarity 100.0%; Pred. No. 0.046;
 Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 6 TTQGVVSALQTQQA 19
 |||||
 DB 57 TTQGVVSALQTQQA 70

RESULT 6
 Q938V1_STRMU
 ID Q938V1_STRMU PRELIMINARY; PRT; 432 AA.
 AC Q938V1;
 DT 01-DEC-2001 (TrEMBLrel. 19, Created)
 DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
 DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
 DE Glucan-binding protein B.
 OS Streptococcus mutans.
 OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
 OC Streptococcus.
 OX NCBI_TaxID=1309;
 RN [1]
 RP NUCLEOTIDE SEQUENCE.
 RC STRAIN=3SN1;
 RA Jin S., Duncan M.J., Taubman M.A., Smith D.J.;
 RT "Cloning of the gbpB gene from Streptococcus mutans.";
 RL J. Dent. Res. 79:224-224(2000).
 RN [2]
 RP NUCLEOTIDE SEQUENCE.
 RC STRAIN=3SN1;
 RX MEDLINE=21481971; PubMed=11598068;
 RX DOI=10.1128/JAI.69.11.6931-6941.2001;
 RA Mattos-Graner R.O., Jin S., King W.F., Chen T., Smith D.J.,
 RA Duncan M.J.;
 RT "Cloning of the Streptococcus mutans gene encoding glucan binding
 RT protein B and analysis of genetic diversity and protein production in
 RT clinical isolates.";
 RL Infect. Immun. 69:6931-6941(2001).
 DR EMBL; AY046413; AAK94503.1; -; Genomic_DNA.
 DR InterPro; IPR007921; CHAP.
 DR InterPro; IPR009148; Siba.
 DR Pfam; PF05257; CHAP; 1.
 DR PRINTS; PR01852; SIBAPROTEIN.
 DR PROSITE; PS50911; CHAP; 1.
 SQ SEQUENCE 432 AA; 44652 MW; 3F88BCB9A1F3BE4F CRC64;

Query Match 67.0%; Score 65; DB 2; Length 432;
 Best Local Similarity 100.0%; Pred. No. 0.046;
 Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 6 TTQGVVSALQTQQA 19
 |||||
 DB 57 TTQGVVSALQTQQA 70
 RESULT 7
 Q938V2_STRMU
 ID Q938V2_STRMU PRELIMINARY; PRT; 432 AA.
 AC Q938V2;
 DT 01-DEC-2001 (TrEMBLrel. 19, Created)
 DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
 DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
 DE Glucan-binding protein B.
 OS Streptococcus mutans.
 OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
 OC Streptococcus.
 OX NCBI_TaxID=1309;
 RN [1]
 RP NUCLEOTIDE SEQUENCE.
 RC STRAIN=15JP2;
 RA Jin S., Duncan M.J., Taubman M.A., Smith D.J.;
 RT "Cloning of the gbpB gene from Streptococcus mutans.";
 RL J. Dent. Res. 79:224-224(2000).
 RN [2]
 RP NUCLEOTIDE SEQUENCE.
 RC STRAIN=15JP2;
 RX MEDLINE=21481971; PubMed=11598068;
 RX DOI=10.1128/JAI.69.11.6931-6941.2001;
 RA Mattos-Graner R.O., Jin S., King W.F., Chen T., Smith D.J.,
 RA Duncan M.J.;
 RT "Cloning of the Streptococcus mutans gene encoding glucan binding
 RT protein B and analysis of genetic diversity and protein production in
 RT clinical isolates.";
 RL Infect. Immun. 69:6931-6941(2001).
 DR EMBL; AY046413; AAK94502.1; -; Genomic_DNA.
 DR InterPro; IPR007921; CHAP.
 DR InterPro; IPR009148; Siba.
 DR Pfam; PF05257; CHAP; 1.
 DR PRINTS; PR01852; SIBAPROTEIN.
 DR PROSITE; PS50911; CHAP; 1.
 SQ SEQUENCE 432 AA; 44648 MW; E769B2504AE50E9 CRC64;

Query Match 67.0%; Score 65; DB 2; Length 432;
 Best Local Similarity 100.0%; Pred. No. 0.046;
 Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 6 TTQGVVSALQTQQA 19
 |||||
 DB 57 TTQGVVSALQTQQA 70

RESULT 8
 Q5KRH3_CORGL
 ID Q5KRH3_CORGL PRELIMINARY; PRT; 130 AA.
 AC Q5KRH3;
 DT 10-MAY-2005 (TrEMBLrel. 30, Created)
 DT 10-MAY-2005 (TrEMBLrel. 30, Last sequence update)
 DT 10-MAY-2005 (TrEMBLrel. 30, Last annotation update)
 DE Hypothetical protein.
 OS Corynebacterium glutamicum (Brevibacterium flavum).
 OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
 OC Corynebacteriaceae; Corynebacteriaceae; Corynebacterium.
 OX NCBI_TaxID=1718;
 RN [1]
 RP NUCLEOTIDE SEQUENCE.
 RC STRAIN=R;
 RA Suzuki N., Okayama S., Nonaka H., Tsuge Y., Inui M., Yukawa H.;
 RT "Large-Scale Engineering of the Corynebacterium glutamicum Genome.";

RL Appl. Environ. Microbiol. 71:3369-3372(2005).
 DR EMBL; AB193035; BAD84114.1; -; Genomic_DNA.
 SQ Hypothetical protein.

Query Match 52.6%; Score 51; DB 2; Length 130;
 Best Local Similarity 58.8%; Pred. No. 3;
 Matches 10; Conservative 4; Mismatches 3; Indels 0; Gaps 0;

QY 4 LITIQGVVSALQTOQAG 20
 Db 102 LVQLNKLKLSALQTOQAG 118

RESULT 9
 Q6ALV1 DESPS PRELIMINARY; PRT; 448 AA.
 AC Q6ALV1
 DT 25-OCT-2004 (TrEMBLrel. 28, Created)
 DT 25-OCT-2004 (TrEMBLrel. 28, Last sequence update)
 DE Probable exodeoxyribonuclease VII, large chain.
 GN OrderedLocusNames=DP1945;
 OS Desulfotalea psychrophila.
 OC Bacteria; Proteobacteria; Deltaproteobacteria; Desulfobacteriales;
 OC Desulfobulbaceae; Desulfotalea.
 OX NCBI_TaxID=84980;
 RN [1]

RP NUCLEOTIDE SEQUENCE.
 RC STRAIN=LSV54 / DSM 12343;
 RX PubMed=15305914; DOI=10.1111/j.1462-2920.2004.00665.x;
 RA Rabus R., Ruepp A., Frickel T., Rattel T., Fartmann B., Stark M.,
 RA Bauer M., Zibat A., Lombardot T., Becker I., Amann J., Gellner K.,
 RA Teeling H., Leuschner W.D., Gloeckner F.-O., Lupas A.N., Amann R.,
 RA Klenk H.-P.;
 RT "The genome of Desulfotalea psychrophila, a sulfate-reducing bacterium
 from permanently cold Arctic sediments.";
 RT Environ. Microbiol. 6:887-902(2004).
 CC -1- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).
 DR EMBL; CR522870; CAG36674.1; -; Genomic DNA.
 DR GO; GO:0009318; C:exodeoxyribonuclease VII complex; IEA.
 DR GO; GO:0008855; F:exodeoxyribonuclease VII activity; IEA.
 DR GO; GO:0004527; F:exonuclease activity; IEA.
 DR GO; GO:0003676; F:nucleic acid binding; IEA.
 DR GO; GO:0006308; P:DNA catabolism; IEA.
 DR InterPro; IPR003439; ABC transp like.
 DR InterPro; IPR003753; Exonuc VII L.
 DR InterPro; IPR004365; OB-Fold trNA.
 DR Pfam; PF02601; Exonuc_VII_L; 1.
 DR Pfam; PF01336; trNA_anti; 1.
 DR TIGRFAMs; TIGR00237; xsea; 1.
 DR PROSITE; PS00211; ABC_TRANSPORTER_1; UNKNOWN_1.
 KW Complete proteome.
 SQ SEQUENCE 448 AA; 50258 MW; B16366C6AB032B48 CRC64;

Query Match 49.5%; Score 48; DB 2; Length 448;
 Best Local Similarity 45.0%; Pred. No. 34;
 Matches 9; Conservative 6; Mismatches 5; Indels 0; Gaps 0;

QY 1 KHKLITIQGVVSALQTOQAG 20
 Db 25 KRFITIRGEISNLKTPYSG 44

RESULT 10
 Q6BUA6 DEBHA PRELIMINARY; PRT; 529 AA.
 ID Q6BUA6 DEBHA
 AC Q6BUA6
 DT 25-OCT-2004 (TrEMBLrel. 28, Created)
 DT 25-OCT-2004 (TrEMBLrel. 28, Last sequence update)
 DT 25-OCT-2004 (TrEMBLrel. 28, Last annotation update)
 DE Similar to CA5768|CAPAC1 Candida albicans.
 GN OrderedLocusNames=DEHA0C13343g;

OS Debaryomyces hansenii (Yeast) (Torulaspora hansenii).
 OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
 OC Saccharomycetales; Saccharomycetaceae; Debaryomycetes.
 OX NCBI_TaxID=4959;
 RN [1]
 RP NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
 RC STRAIN=ATCC 36239 / CBS 767;
 RX PubMed=15229592; DOI=10.1038/nature02579;
 RA DuJon B., Sherman D., Fischer G., Durrens P., Casaregola S.,
 RA Lafontaine I., de Montigny J., Marck C., Neuveglise C., Falla E.,
 RA Goffard N., Frangeul L., Aigle M., Anthouard V., Babour A., Barbe V.,
 RA Barnay S., Blanchin S., Beckerich J.-M., Beyne E., Bleykasten C.,
 RA Boisarame A., Boyer J., Cattolico L., Confaniolieri F., de Daruvar A.,
 RA Despons L., Fabre E., Fairhead C., Ferry-Dumazet H., Groppi A.,
 RA Hantraye F., Hennequin C., Jauniaux N., Joyet P., Kachouri R.,
 RA Kerrest A., Koszul R., Lemaire M., Lesur I., Ma L., Muller R.,
 RA Nicaud J.-M., Nikolaki M., Oztas S., Ozier-Kalogeropoulos O.,
 RA Pellenz S., Potier S., Richard G.-F., Straub M.-L., Suleau A.,
 RA Swennen D., Tekala F., Wesolowski-Louvel M., Westhof E., Wirth B.,
 RA Zenlou-Meyer M., Zivanovic Y., Bolotin-Fukuhara M., Thierry A.,
 RA Bouchier C., Caudron B., Scarpelli C., Gaillardin C., Weissenbach J.,
 RA Wincker P., Souciet J.-L.;
 RT "Genome evolution in yeasts.";
 RL Nature 430:35-44(2004).
 DR EMBL; CR382135; CAG86289.1; -; Genomic_DNA.
 DR InterPro; IPR001680; WD40.
 DR Pfam; PF00400; WD40; 7.
 DR PRINTS; PR00320; GPROTEINBRPT.
 DR ProDom; PD000018; WD40; 2.
 DR SMART; SM00320; WD40; 7.
 DR PROSITE; PS00678; WD_REPEATS_1; 4.
 DR PROSITE; PS00682; WD_REPEATS_2; 3.
 DR PROSITE; PS00294; WD_REPEATS_REGION; 1.
 KW Complete proteome; Repeat; WD repeat.
 SQ SEQUENCE 529 AA; 58922 MW; A96740E457515E0E CRC64;

Query Match 48.5%; Score 47; DB 2; Length 529;
 Best Local Similarity 50.0%; Pred. No. 59;

Matches 8; Conservative 6; Mismatches 2; Indels 0; Gaps 0;

QY 1 KHKLITIQGVVSALQOT 16
 Db 70 QHKIIDLEGEVSNLAT 85

RESULT 11
 Q4T8Q5 TETNG PRELIMINARY; PRT; 223 AA.
 ID Q4T8Q5 TETNG
 AC Q4T8Q5
 DT 13-SEP-2005 (TrEMBLrel. 31, Created)
 DT 13-SEP-2005 (TrEMBLrel. 31, Last sequence update)
 DT 13-SEP-2005 (TrEMBLrel. 31, Last annotation update)
 DE Chromosome undetermined SCAF7749, whole genome shotgun sequence.
 DE (Fragment).
 GN ORFNames=GSTENG0005121001;
 OS Tetraodon nigroviridis (Green puffer).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;
 OC Acanthomorpha; Acanthopterygii; Percomorpha; Tetraodontiformes;
 OC Tetraodontidae; Tetraodontidae; Tetraodon.
 OX NCBI_TaxID=99883;
 RN [1]

RP NUCLEOTIDE SEQUENCE.
 RA Jaillon O., Aury J.M., Brunet F., Petit J.L., Stange-Thomann N.,
 RA Mauceli E., Bouneau L., Fischer C., Ozouf-Costaz C., Bernot A.,
 RA Nicaud S., Jaffe D., Fisher S., Lutfalla G., Dossat C., Segurens B.,
 RA Dasilva C., Salancoubat M., Levy M., Boudet N., Castellano S.,
 RA Anthouard V., Jubin C., Cattolico L., Katinka M., Vacherie B.,
 RA Blemond C., Skalli Z., Cattolico L., Poulain J., De Berardinis V.,
 RA Craud C., Duprat G., Brottier P., Coutanceau J.P., Gouzy J.,
 RA Farra G., Lardier S., Chapelle C., McKernan K.J., McSwan P., Bosak S.,
 RA Kellis M., Volff J.N., Guigo R., Zody M.C., Mesirov J.,
 RA Lindblad-Toh K., Birren B., Nusbaum C., Kahn D., Robinson-Rechavi M.,

```

RA Laudet V., Schachter V., Quetier F., Saurin W., Scarpelli C.,
RT Wincker P., Iander E.S., Weissenbach J., Roest Crolius H.;
RT "Genome duplication in the teleost fish Tetraodon nigroviridis reveals
RL the early vertebrate proto-karyotype.";
RL Nature 431:946-957(2004).
RN [2]
RP NUCLEOTIDE SEQUENCE
RG Genoscope; Whitehead Institute Centre for Genome Research;
RG Submitted (FEB-2004) to the EMBL/GenBank/DBJ databases.
CC -1- CAUTION: The sequence shown here is derived from an
CC EMBL/GenBank/DBJ whole genome shotgun (WGS) entry which is
CC preliminary data.
DR EMBL: CAEE01007749; CAF90727.1; -; Genomic_DNA.
FT NON_TER 1 223
FT NON_TER 223 223
SQ SEQUENCE 223 AA; 24595 MW; 0FF76D0E6AAEEDA CRC64;

Query Match 47.4%; Score 46; DB 2; Length 223;
Best Local Similarity 36.8%; Pred. No. 36;
Matches 7; Conservative 7; Mismatches 5; Indels 0; Gaps 0;

QY 2 HKLITIQGVSLQTOOAG 20
   |||:||||:|:|:|
Db 60 HKLLVLEGQTSLSRCRAVG 78

RESULT 12
ID Q82LB1 STRAW PRELIMINARY; PRT; 465 AA.
AC Q82LB1
DT 01-JUN-2003 (TrEMBLrel. 24, Created)
DT 01-JUN-2003 (TrEMBLrel. 24, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Hypothetical protein.
OS OrderedLocusNames=SAV2100;
GN Streptomyces avermitilis.
OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
OC Streptomycineae; Streptomycetaceae; Streptomyces.
OX NCBI_TaxID=33903;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=MA-4680 / ATCC 31267 / NCIMB 12804 / NRRL 8165;
RC MEDLINE=22608306; PubMed=12692562; DOI=10.1038/nbt820;
RX Ikeda H., Ishikawa J., Hanamoto A., Shinose M., Kikuchi T.,
RX Sakaki Y., Hattori M., Omura S.;
RA "Complete genome sequence and comparative analysis of the industrial
RT microorganism Streptomyces avermitilis.";
RL Nat. Biotechnol. 21:526-531(2003).
RN [2]
RP NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
RC STRAIN=MA-4680 / ATCC 31267 / NCIMB 12804 / NRRL 8165;
RC MEDLINE=21477403; PubMed=11572948; DOI=10.1073/pnas.211433198;
RX Omura S., Ikeda H., Ishikawa J., Hanamoto A., Takahashi C.,
RA Shinose M., Takahashi Y., Horikawa H., Nakazawa H., Osonoe T.,
RA Kikuchi H., Shiba T., Sakaki Y., Hattori M.;
RT "Genome sequence of an industrial microorganism Streptomyces
RT avermitilis: deducing the ability of producing secondary
RT metabolites.";
RL Proc. Natl. Acad. Sci. U.S.A. 98:12215-12220(2001).
DR EMBL: BA000030; BAC69811.1; -; Genomic_DNA.
KW Complete proteome; Hypothetical protein.
SQ SEQUENCE 465 AA; 49845 MW; F6B451DEED3170A5 CRC64;

Query Match 47.4%; Score 46; DB 2; Length 465;
Best Local Similarity 44.4%; Pred. No. 76;
Matches 8; Conservative 5; Mismatches 5; Indels 0; Gaps 0;

QY 2 HKLITIQGVSLQTOOQA 19
   |||:||||:|:|:|
Db 297 HRLVPVHGRPSSLHTAQ 314

RESULT 13

```

```

Q59UT0 CANAL PRELIMINARY; PRT; 592 AA.
ID Q59UT0
AC Q59UT0
DT 10-MAY-2005 (TrEMBLrel. 30, Created)
DT 10-MAY-2005 (TrEMBLrel. 30, Last sequence update)
DT 10-MAY-2005 (TrEMBLrel. 30, Last annotation update)
DE Potential peroxisome targeting sequence receptor Pex5.
GN Names=Pex5; ORFNames=CAO19.13085, CAO19.5640;
OS Candida albicans SC5314.
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC Saccharomycetales; mitosporic Saccharomycetales; Candida.
OX NCBI_TaxID=237561;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=SC5314;
RC PubMed=15123810; DOI=10.1073/pnas.0401648101;
RA Jones T., Federspiel N.A., Chibana H., Dungan J., Kalman S.,
RA Magee B.B., Newport G., Thorstenson Y.R., Agabian N., Magee P.T.,
RA Davis R.W., Scherer S.;
RA "The diploid genome sequence of Candida albicans.";
RT Proc. Natl. Acad. Sci. U.S.A. 101:7329-7334(2004).
RN [2]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=SC5314;
RA Dungan J., Kuo A., Newport G., Lan C.-Y., Iijima C., Adegbola O.,
RA Roberts J., Persson K., Donnelly S., Favoreto S., Tzung K.-W.,
RA Jones T., Scherer S., Agabian N.;
RT "Annotation of the Genome of Candida albicans.";
RL Submitted (APR-2004) to the EMBL/GenBank/DBJ databases.
DR EMBL: AACQ01000129; EAK94204.1; -; Genomic_DNA.
DR EMBL: AACQ01000128; EAK94251.1; -; Genomic_DNA.
DR GO: GO:0004872; Fireceptor activity; IEA.
DR InterPro: IPR001484; Pyrokinin.
DR InterPro: IPR001440; TPR.
DR InterPro: IPR011990; TPR-like_helical.
DR Pfam: PF00515; TPR_1; 4.
DR SMART: SM00028; TPR; 4.
DR PROSITE: PS00539; PYROKININ; UNKNOWN_1.
DR PROSITE: PS50005; TPR; 4.
DR PROSITE: PS50293; TPR_REGION; 1.
KW Receptor; Repeat; TPR_repeat.
SQ SEQUENCE 592 AA; 67324 MW; FA5B271F220B38E6 CRC64;

Query Match 47.4%; Score 46; DB 2; Length 592;
Best Local Similarity 47.4%; Pred. No. 97;
Matches 9; Conservative 5; Mismatches 5; Indels 0; Gaps 0;

QY 1 HKLITIQGVSLQTOOQA 19
   |||:||||:|:|:|
Db 80 RHELTIONQPNAIHQOQS 98

RESULT 14
Q9LPZ2 ARATH PRELIMINARY; PRT; 636 AA.
ID Q9LPZ2
AC Q9LPZ2
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE T23J18.9.
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;
OC eurosid II; Brassicales; Brassicaceae; Arabidopses.
OX NCBI_TaxID=3702;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RA Shinn P., Brooks S., Buehler E., Chao Q., Johnson-Hopson C., Khan S.,
RA Kim C., Altafi H., Bei Q., Chin C., Chiu J., Choi E., Conn L.,
RA Conway A., Gonzales A., Hansen N., Howing B., Koo T., Lam B., Lee J.,
RA Lenz C., Li J., Liu A., Liu K., Liu S., Mukharsky N., Nguyen M.,
RA Palm C., Pham P., Sakano H., Schwartz J., Southwick A., Thaveri A.,
RA Toriumi M., Vaysberg M., Yu G., Federspiel N.A., Theologis A.,

```


GenCore version 5.1.6
Copyright (c) 1993 - 2006 Compugen Ltd.

OM protein - protein search, using sw model

Run on: January 25, 2006, 18:50:05 ; Search time 16.6 Seconds
(without alignments)
99.609 Million cell updates/sec

Title: US-10-797-821-2

Perfect score: 97

Sequence: 1 KKLITIQQVSALQTOQAG 20

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 572060 seqs, 82675679 residues

Total number of hits satisfying chosen parameters: 572060

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

- Issued Patents AA:*
- 1: /cgn2_6/prodata/1/iaa/5_COMB.pep.*
 - 2: /cgn2_6/prodata/1/iaa/6_COMB.pep.*
 - 3: /cgn2_6/prodata/1/iaa/H_COMB.pep.*
 - 4: /cgn2_6/prodata/1/iaa/PCUTS_COMB.pep.*
 - 5: /cgn2_6/prodata/1/iaa/RE_COMB.pep.*
 - 6: /cgn2_6/prodata/1/iaa/backfilese1.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	44	45.4	295	2	US-09-134-000C-6619
2	44	45.4	410	2	US-09-205-448-1
3	44	45.4	789	2	US-09-831-846-2
4	43	44.3	371	2	US-09-270-767-42554
5	42.5	43.8	700	2	US-09-252-991A-32447
6	42	43.3	192	2	US-09-583-110-5295
7	42	43.3	192	2	US-09-107-433-4762
8	42	43.3	317	1	US-08-726-306A-168
9	42	43.3	1020	2	US-09-538-092-911
10	41	42.3	292	2	US-09-134-001C-3479
11	41	42.3	1024	2	US-09-562-737-45
12	40	41.2	430	2	US-09-489-039A-7776
13	40	41.2	536	2	US-09-902-540-13091
14	40	41.2	720	2	US-09-394-272-14
15	40	41.2	1267	2	US-09-543-681A-6130
16	40	41.2	1433	2	US-09-487-558B-60
17	39	40.2	75	2	US-09-270-767-56996
18	39	40.2	207	2	US-10-104-047-2278
19	39	40.2	210	2	US-09-222-938A-67
20	39	40.2	289	2	US-09-252-991A-19366
21	39	40.2	378	2	US-09-252-991A-20726
22	39	40.2	392	2	US-09-583-110-4374
23	39	40.2	399	2	US-09-107-433-3230
24	39	40.2	443	2	US-09-252-991A-20035
25	39	40.2	641	2	US-09-583-110-4263
26	39	40.2	641	2	US-09-107-433-2948
27	39	40.2	675	2	US-09-171-878-1

28	39	40.2	725	2	US-09-270-767-41750	Sequence 41750, A
29	39	40.2	1013	2	US-09-712-363-214	Sequence 214, App
30	39	40.2	1024	2	US-09-562-737-43	Sequence 43, Appl
31	38.5	39.7	118	2	US-09-489-039A-12533	Sequence 12533, A
32	38.5	39.7	661	2	US-09-252-991A-32899	Sequence 32899, A
33	38	39.2	87	2	US-09-270-767-61632	Sequence 61632, A
34	38	39.2	250	1	US-08-872-961A-1	Sequence 1, Appli
35	38	39.2	250	2	US-09-231-258-1	Sequence 1, Appli
36	38	39.2	274	2	US-09-270-767-46079	Sequence 46079, A
37	38	39.2	307	2	US-09-270-767-41998	Sequence 41998, A
38	38	39.2	612	2	US-09-303-518D-44	Sequence 44, Appl
39	38	39.2	692	2	US-09-252-991A-25793	Sequence 25793, A
40	38	39.2	701	2	US-09-252-991A-26940	Sequence 26940, A
41	38	39.2	747	1	US-08-816-693A-51	Sequence 51, Appl
42	38	39.2	747	2	US-08-885-291-51	Sequence 51, Appl
43	38	39.2	747	2	US-09-496-672-51	Sequence 51, Appl
44	38	39.2	1024	2	US-09-562-737-41	Sequence 41, Appl
45	37	38.1	54	2	US-09-434-840-76	Sequence 76, Appl

ALIGNMENTS

RESULT 1

US-09-134-000C-6619
; Sequence 6619, Application US/09134000C
; Patent No. 6617156
; GENERAL INFORMATION:
; APPLICANT: Lynn Doucette-Stamm et al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO
; FILE REFERENCE: 032796-032
; CURRENT APPLICATION NUMBER: US/09/134,000C
; CURRENT FILING DATE: 1998-08-13
; PRIOR APPLICATION NUMBER: US 60/055,778
; PRIOR FILING DATE: 1997-08-15
; NUMBER OF SEQ ID NOS: 6812
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 6619
; LENGTH: 295
; TYPE: PRT
; ORGANISM: Enterococcus faecalis
US-09-134-000C-6619

Query Match 45.4%; Score 44; DB 2; Length 295;
Best Local Similarity 45.0%; Pred. No. 9.6;
Matches 9; Conservative 3; Mismatches 8; Indels 0; Gaps 0;
QY 1 KKLITIQQVSALQTOQAG 20
DB 227 KFLVTRQKEAINTDPNG 246

RESULT 2

US-09-205-448-1
; Sequence 1, Application US/09205448
; Patent No. 6537746
; GENERAL INFORMATION:
; APPLICANT: Arnold, Frances
; APPLICANT: Shao, Zhixin
; APPLICANT: Volkov, Alexander
; TITLE OF INVENTION: Method for Creating Polynucleotide and Polypeptide
; FILE REFERENCE: 018097-025710PC
; CURRENT APPLICATION NUMBER: US/09/205,448
; CURRENT FILING DATE: 1998-12-04
; PRIOR APPLICATION NUMBER: US 60/067,908
; PRIOR FILING DATE: 1997-12-08
; NUMBER OF SEQ ID NOS: 11
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 1
; LENGTH: 410

```
; TYPE: PRT
; ORGANISM: Rhizobium lupini
; FEATURE:
; OTHER INFORMATION: flagellin A (Flaa)
US-09-205-448-1

Query Match      45.4%; Score 44; DB 2; Length 410;
Best Local Similarity 40.0%; Pred. No. 14;
Matches 12; Conservative 3; Mismatches 5; Indels 10; Gaps 1;

QY 1 KHKLIT-----IQGVLSALQTQQAG 20
Db 92 KKKLVTAQESSADTKYIQGVKQLQELKG 121

RESULT 3
US-09-831-846-2
; Sequence 2, Application US/09831846
; Patent No. 6852836
; GENERAL INFORMATION:
; APPLICANT: Funahashi, Shin-ichi
; APPLICANT: Miyata, Shoji
; APPLICANT: No. 6852836ura. No. 6852836uo
; APPLICANT: Nagase, Takahiro
; APPLICANT: Ohara, Osamu
; TITLE OF INVENTION: NOVEL GENE ENCODING BRAIN-SPECIFIC MEMBRANE PROTEIN
; FILE REFERENCE: 06501-081001
; CURRENT APPLICATION NUMBER: US/09/831,846
; CURRENT FILING DATE: 2001-08-21
; PRIOR APPLICATION NUMBER: PCT/JP99/06449
; PRIOR FILING DATE: 1999-11-18
; PRIOR APPLICATION NUMBER: JP 10/331727
; PRIOR FILING DATE: 1998-11-20
; NUMBER OF SEQ ID NOS: 7
; SEQ ID NO 2
; LENGTH: 789
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-831-846-2

Query Match      45.4%; Score 44; DB 2; Length 789;
Best Local Similarity 36.8%; Pred. No. 31;
Matches 7; Conservative 6; Mismatches 6; Indels 6; Gaps 0;

QY 2 HKLITIQGVLSALQTQQAG 20
Db 296 HKLLVLEGGQATLKCKAIG 314

RESULT 4
US-09-270-767-42554
; Sequence 42554, Application US/09270767
; Patent No. 6703491
; GENERAL INFORMATION:
; APPLICANT: Homburger et al.
; TITLE OF INVENTION: Nucleic acids and proteins of Drosophila melanogaster
; FILE REFERENCE: File Reference: 7326-094
; CURRENT APPLICATION NUMBER: US/09/270,767
; CURRENT FILING DATE: 1999-03-17
; NUMBER OF SEQ ID NOS: 62517
; SOFTWARE: Patent in Ver. 2.0
; SEQ ID NO 42554
; LENGTH: 371
; TYPE: PRT
; ORGANISM: Drosophila melanogaster
; FEATURE:
; OTHER INFORMATION: Xaa means any amino acid
US-09-270-767-42554

Query Match      44.3%; Score 43; DB 2; Length 371;
Best Local Similarity 56.2%; Pred. No. 19;
Matches 9; Conservative 4; Mismatches 3; Indels 3; Gaps 0;
```

```
QY 3 KLITIQGVLSALQTQQ 18
Db 134 KLISLGLQCAICTQQ 149
```

```
RESULT 5
US-09-252-991A-32447
; Sequence 32447, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 32447
; LENGTH: 700
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-32447
```

```
Query Match      43.8%; Score 42.5; DB 2; Length 700;
Best Local Similarity 45.0%; Pred. No. 50;
Matches 9; Conservative 5; Mismatches 5; Indels 1; Gaps 1;
```

```
QY 2 HKLITIQGVLSALQTQQAG 20
Db 147 HELVEVAGDQQQAVQARQAG 166
```

```
RESULT 6
US-09-583-110-5295
; Sequence 5295, Application US/09583110
; Patent No. 6699703
; GENERAL INFORMATION:
; APPLICANT: Lynn Doucette-Stamm et al.
; TITLE OF INVENTION: Nucleic Acid and Amino Acid Sequences Relating to Streptococcus
; TITLE OF INVENTION: Pneumoniae for Diagnostics and Therapeutics
; FILE REFERENCE: PATH00-07A
; CURRENT APPLICATION NUMBER: US/09/583,110
; CURRENT FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: US 09/107,433
; PRIOR FILING DATE: 1998-06-30
; PRIOR APPLICATION NUMBER: US 60/085,131
; PRIOR FILING DATE: 1998-05-12
; PRIOR APPLICATION NUMBER: US 60/051,553
; PRIOR FILING DATE: 1997-07-02
; NUMBER OF SEQ ID NOS: 5322
; SEQ ID NO 5295
; LENGTH: 192
; TYPE: PRT
; ORGANISM: Streptococcus pneumoniae
US-09-583-110-5295
```

```
Query Match      43.3%; Score 42; DB 2; Length 192;
Best Local Similarity 56.2%; Pred. No. 13;
Matches 9; Conservative 2; Mismatches 5; Indels 0; Gaps 0;
```

```
QY 5 ITIQGVLSALQTQQAG 20
Db 37 INIIGQAPGLKTQAG 52
```

```
RESULT 7
US-09-107-433-4762
; Sequence 4762, Application US/09107433
; Patent No. 6800744
```

```

; GENERAL INFORMATION:
; APPLICANT: Lynn A Doucette-Stamm and David Bush
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID
; SEQUENCES RELATING TO STREPTOCOCCUS PNEUMONIAE FOR DIAGNOSTIC THERAPEUTICS
;
; NUMBER OF SEQUENCES: 5206
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: GENOME THERAPEUTICS CORPORATION
; STREET: 100 Beaver Street
; CITY: Waltham
; STATE: Massachusetts
; COUNTRY: USA
;
; ZIP: 02354
;
; COMPUTER READABLE FORM:
; MEDIUM TYPE: CD-ROM ISO9660
; COMPUTER: <Unknown>
; OPERATING SYSTEM: <Unknown>
; SOFTWARE: <Unknown>
;
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/107,433
; FILING DATE: 30-Jun-1998
;
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 60/ 085131
; FILING DATE: May 12, 1998
; APPLICATION NUMBER: 60/051553
; FILING DATE: July 2, 1997
;
; ATTORNEY/AGENT INFORMATION:
; NAME: Ariniello, Pamela Deneke
; REGISTRATION NUMBER: 40,489
; REFERENCE/DOCKET NUMBER: GTC-011
;
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (781)893-5007
; TELEFAX: (781)893-8277
;
; INFORMATION FOR SEQ ID NO: 4762:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 192 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; HYPOTHETICAL: YES
; ORIGINAL SOURCE:
; ORGANISM: Streptococcus pneumoniae
;
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (B) LOCATION 1...192
; SEQUENCE DESCRIPTION: SEQ ID NO: 4762:
US-09-107-433-4762

Query Match 43.3%; Score 42; DB 2; Length 192;
Best Local Similarity 56.2%; Pred. No. 13;
Matches 9; Conservative 2; Mismatches 5; Indels 0; Gaps 0;

QY 5 ITIQGVVSALQTQAG 20
Db 37 INIQAGPKLTQAG 52

RESULT 8
US-08-726-306A-168
; Sequence 168, Application US/08726306A
; Patent No. 5958684
; GENERAL INFORMATION:
; APPLICANT: van Leeuwen, Frederik Willem
; APPLICANT: Burbach, Johannes Peter Henri
; APPLICANT: Grosveld, Franklin G.
; TITLE OF INVENTION: DIAGNOSIS METHOD AND REAGENTS
; NUMBER OF SEQUENCES: 189
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Banner & Witcoff, Ltd.
; STREET: 1 Financial Center
; CITY: Boston
; STATE: MA
; COUNTRY: US

; GENERAL INFORMATION:
; APPLICANT: Lynn A Doucette-Stamm and David Bush
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID
; SEQUENCES RELATING TO STREPTOCOCCUS PNEUMONIAE FOR DIAGNOSTIC THERAPEUTICS
;
; NUMBER OF SEQUENCES: 5206
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: GENOME THERAPEUTICS CORPORATION
; STREET: 100 Beaver Street
; CITY: Waltham
; STATE: Massachusetts
; COUNTRY: USA
;
; ZIP: 02354
;
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.50 inch, 1.44 Mb storage
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: WordPerfect 6.1
;
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/726,306A
; FILING DATE: 02-Oct-1996
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: GB 95/20080.4
; FILING DATE: 02-Oct-1995
;
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 60/009,832
; FILING DATE: 01-Jan-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: Williams, Ph.D., Kathleen M.
; REGISTRATION NUMBER: 34,380
; REFERENCE/DOCKET NUMBER: 96,048-A (3255/00784)
;
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (617) 345-9100
; TELEFAX: (617) 345-9111
;
; INFORMATION FOR SEQ ID NO: 168:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 317 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: unknown
; MOLECULE TYPE: peptide
;
; US-08-726-306A-168

Query Match 43.3%; Score 42; DB 1; Length 317;
Best Local Similarity 46.7%; Pred. No. 24;
Matches 7; Conservative 6; Mismatches 2; Indels 0; Gaps 0;

QY 6 TIQGVVSALQTQAG 20
Db 147 SLIEGAALURQQAG 161

RESULT 9
US-09-538-092-911
; Sequence 911, Application US/09538092
; Patent No. 6753314
; GENERAL INFORMATION:
; APPLICANT: Giot, Loic
; APPLICANT: Mansfield, Traci A.
; TITLE OF INVENTION: Protein-Protein Complexes and Method of Using Same
; FILE REFERENCE: 15966-542
; CURRENT APPLICATION NUMBER: US/09/538,092
; CURRENT FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: 60/127,352
; PRIOR FILING DATE: 1999-04-01
; PRIOR APPLICATION NUMBER: 60/178,965
; PRIOR FILING DATE: 2000-02-01
; NUMBER OF SEQ ID NOS: 1387
; SOFTWARE: CuraPatSeqFormatter Version 0.9
; SEQ ID NO 911
; LENGTH: 1020
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (0)...(0)
; OTHER INFORMATION: Polypeptide Accession Number P12036
US-09-538-092-911

Query Match 43.3%; Score 42; DB 2; Length 1020;
Best Local Similarity 46.7%; Pred. No. 97;
Matches 7; Conservative 6; Mismatches 2; Indels 0; Gaps 0;

QY 6 TIQGVVSALQTQAG 20
::: :|||
```

```

Db      124 SLEGEAALRQQAG 138

RESULT 10
US-09-134-001C-3479
; Sequence 3479, Application US/09134001C
; Patent No. 6380370
; GENERAL INFORMATION:
; APPLICANT: Lynn Doucette-Stamm et al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO STAPHYLOCOCCUS
; TITLE OF INVENTION: EPIDERMIDIS FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: GTC-007
; CURRENT APPLICATION NUMBER: US/09/134,001C
; CURRENT FILING DATE: 1998-08-13
; PRIOR APPLICATION NUMBER: US 60/064,964
; PRIOR FILING DATE: 1997-11-08
; PRIOR APPLICATION NUMBER: US 60/055,779
; PRIOR FILING DATE: 1997-08-14
; NUMBER OF SEQ ID NOS: 5674
; SEQ ID NO 3479
; LENGTH: 292
; TYPE: PRT
; ORGANISM: Staphylococcus epidermidis
US-09-134-001C-3479

Query Match      42.3%; Score 41; DB 2; Length 292;
Best Local Similarity 64.3%; Pred. No. 33;
Matches 9; Conservative 2; Mismatches 2; Indels 3; Gaps 0;

QY      6 TTQGVSAALQTQQA 19
      ||:||||:|
Db      269 TTQGVSTVATGQA 282

RESULT 11
US-09-562-737-45
; Sequence 45, Application US/09562737
; Patent No. 6428967
; GENERAL INFORMATION:
; APPLICANT: Herz, Joachim
; APPLICANT: Gotthardt, Michael
; TITLE OF INVENTION: LDL Receptor Signaling Pathways
; FILE REFERENCE: UTSW0708
; CURRENT APPLICATION NUMBER: US/09/562,737
; CURRENT FILING DATE: 2000-05-01
; NUMBER OF SEQ ID NOS: 132
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 45
; LENGTH: 1024
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Synthetic
; OTHER INFORMATION: Sequence
US-09-562-737-45

Query Match      42.3%; Score 41; DB 2; Length 1024;
Best Local Similarity 69.2%; Pred. No. 1.5e+02;
Matches 9; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY      6 TTQGVSAALQTQQA 18
      ||| |||||
Db      1010 TTQGVSAWLSQ 1022

RESULT 12
US-09-489-039A-7776
; Sequence 7776, Application US/09489039A
; Patent No. 6610836
; GENERAL INFORMATION:
; APPLICANT: Gary Breton et. al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO KLEBSIELLA
; TITLE OF INVENTION: PNEUMONIAE FOR DIAGNOSTICS AND THERAPEUTICS

```

```

; FILE REFERENCE: 2709.2004001
; CURRENT APPLICATION NUMBER: US/09/489,039A
; CURRENT FILING DATE: 2000-01-27
; PRIOR APPLICATION NUMBER: US 60/117,747
; PRIOR FILING DATE: 1999-01-29
; NUMBER OF SEQ ID NOS: 14342
; SEQ ID NO 7776
; LENGTH: 430
; TYPE: PRT
; ORGANISM: Klebsiella pneumoniae
US-09-489-039A-7776

Query Match      41.2%; Score 40; DB 2; Length 430;
Best Local Similarity 40.0%; Pred. No. 79;
Matches 6; Conservative 5; Mismatches 4; Indels 0; Gaps 0;

QY      1 KHKLITIQGVSAALQ 15
      ||:||||:|
Db      401 KHEFVTLGMEKAVQ 415

RESULT 13
US-09-902-540-13091
; Sequence 13091, Application US/09902540
; Patent No. 6833447
; GENERAL INFORMATION:
; APPLICANT: Goldman, Barry S.
; APPLICANT: Hinkle, Gregory J.
; APPLICANT: Slater, Steven C.
; APPLICANT: Wiegand, Roger C.
; TITLE OF INVENTION: Myxococcus xanthus Genome Sequences and Uses Thereof
; FILE REFERENCE: 38-10(15849)B
; CURRENT APPLICATION NUMBER: US/09/902,540
; CURRENT FILING DATE: 2001-07-10
; PRIOR APPLICATION NUMBER: 60/217,883
; PRIOR FILING DATE: 2000-07-10
; NUMBER OF SEQ ID NOS: 16825
; SEQ ID NO 13091
; LENGTH: 536
; TYPE: PRT
; ORGANISM: Myxococcus xanthus
US-09-902-540-13091

Query Match      41.2%; Score 40; DB 2; Length 536;
Best Local Similarity 47.1%; Pred. No. 1e+02;
Matches 8; Conservative 3; Mismatches 6; Indels 0; Gaps 0;

QY      4 LITIQGVSAALQTQQA 20
      :||:|||||
Db      58 IATLSERVQLQQAG 74

RESULT 14
US-09-394-272-14
; Sequence 14, Application US/09394272
; Patent No. 6472588
; GENERAL INFORMATION:
; APPLICANT: Haigler, Candace H.
; APPLICANT: Holaday, A. Scott
; TITLE OF INVENTION: TRANSGENIC FIBER PRODUCING PLANTS WITH INCREASED
; TITLE OF INVENTION: EXPRESSION OF SUCROSE PHOSPHATE SYNTHASE
; FILE REFERENCE: 201304/1000
; CURRENT APPLICATION NUMBER: US/09/394,272
; CURRENT FILING DATE: 1999-09-10
; NUMBER OF SEQ ID NOS: 14
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 14
; LENGTH: 720
; TYPE: PRT
; ORGANISM: Synechocystis sp.
US-09-394-272-14

Query Match      41.2%; Score 40; DB 2; Length 720;

```


Best Local Similarity 56.2%; Pred. No. 1.5e+02;
Matches 9; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

QY 2 HKLITIQGVSAQTQ 17
||| |
Db 269 HKLIAAYGQSPQLQAQ 284

RESULT 15
US-09-543-681A-6130
; Sequence 6130, Application US/09543681A
; Patent No. 6605709
; GENERAL INFORMATION:
; APPLICANT: GARY BRETON
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PROTEUS MIRABILIS
; FILE REFERENCE: 2709.1002-001
; CURRENT APPLICATION NUMBER: US/09/543,681A
; CURRENT FILING DATE: 2000-04-05
; PRIOR APPLICATION NUMBER: US 60/128,706
; PRIOR FILING DATE: 1999-04-09
; NUMBER OF SEQ ID NOS: 8344
; SEQ ID NO 6130
; LENGTH: 1267
; TYPE: PRT
; ORGANISM: Proteus mirabilis
US-09-543-681A-6130

Query Match 41.2%; Score 40; DB 2; Length 1267;
Best Local Similarity 53.3%; Pred. No. 2.9e+02;
Matches 8; Conservative 4; Mismatches 3; Indels 0; Gaps 0;

QY 5 ITIQGVSAQTQQA 19
: : : : :
Db 784 VDLTGQVNAQSOLA 798

Search completed: January 25, 2006, 19:18:12
Job time : 17.6 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2006 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: January 25, 2006, 19:13:51 ; Search time 60 Seconds
(without alignments)
139.276 Million cell updates/sec

Title: US-10-797-821-2
Perfect score: 97
Sequence: 1 KHLITIQGVSAALQTQQAG 20

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1867569 seqs, 417829326 residues

Total number of hits satisfying chosen parameters: 1867569

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications AA Main:
1: /cgn2_6/prodata/1/pubpaa/US07_PUBCOMB.pap.*
2: /cgn2_6/prodata/1/pubpaa/US08_PUBCOMB.pap.*
3: /cgn2_6/prodata/1/pubpaa/US09_PUBCOMB.pap.*
4: /cgn2_6/prodata/1/pubpaa/US10A_PUBCOMB.pap.*
5: /cgn2_6/prodata/1/pubpaa/US10B_PUBCOMB.pap.*
6: /cgn2_6/prodata/1/pubpaa/US11_PUBCOMB.pap.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Match	Query Length	DB ID	Description
1	97	100.0	20	4	US-10-383-930-2
2	97	100.0	20	5	US-10-797-821-2
3	65	67.0	20	4	US-10-383-930-9
4	65	67.0	20	5	US-10-797-821-9
5	65	67.0	431	4	US-10-383-930-29
6	65	67.0	431	4	US-10-383-930-30
7	65	67.0	431	4	US-10-383-930-33
8	65	67.0	431	5	US-10-797-821-29
9	65	67.0	431	5	US-10-797-821-30
10	65	67.0	431	5	US-10-797-821-33
11	65	67.0	432	4	US-10-383-930-31
12	65	67.0	432	4	US-10-383-930-32
13	65	67.0	432	5	US-10-797-821-31
14	65	67.0	432	5	US-10-797-821-32
15	51	52.6	20	4	US-10-383-930-8
16	51	52.6	20	5	US-10-797-821-8
17	48	49.5	173	4	US-10-425-115-338496
18	46	47.4	465	4	US-10-156-761-9638
19	46	47.4	604	4	US-10-171-404A-6
20	45	46.4	1034	4	US-10-282-122A-75385
21	44	45.4	410	3	US-09-205-448-1
22	44	45.4	410	4	US-10-125-692-31
23	44	45.4	410	4	US-10-371-168-1
24	44	45.4	410	5	US-10-991-347-31
25	44	45.4	412	4	US-10-282-122A-66950
26	44	45.4	456	4	US-10-238-075-314
27	44	45.4	788	4	US-10-004-378A-43

28	44	45.4	789	4	US-10-004-378A-42	Sequence 42, Appl
29	44	45.4	789	4	US-10-291-172-261	Sequence 261, App
30	44	45.4	789	4	US-10-221-278-261	Sequence 261, App
31	44	45.4	789	6	US-11-007-517-2	Sequence 2, Appli
32	44	45.4	832	3	US-09-732-436-24	Sequence 24, Appli
33	44	45.4	832	4	US-10-004-378A-41	Sequence 41, Appl
34	44	45.4	834	4	US-10-425-114-55846	Sequence 55846, A
35	44	45.4	834	4	US-10-425-115-309227	Sequence 309227, A
36	44	45.4	851	4	US-10-369-493-13369	Sequence 13369, A
37	43	44.3	872	4	US-10-047-260-38	Sequence 38, Appl
38	43	44.3	872	4	US-10-310-154-484	Sequence 484, App
39	43	44.3	872	5	US-10-369-493-2762	Sequence 2762, Ap
40	43	44.3	872	5	US-10-732-923-426	Sequence 426, App
41	43	44.3	872	5	US-10-732-923-6939	Sequence 6939, Ap
42	42.5	43.8	418	4	US-10-425-114-48565	Sequence 48565, A
43	42.5	43.8	515	4	US-10-425-115-352245	Sequence 352245, A
44	42.5	43.8	964	5	US-10-473-127-424	Sequence 424, App
45	42.5	43.8	964	5	US-10-473-127-428	Sequence 428, App

ALIGNMENTS

RESULT 1
US-10-383-930-2
; Sequence 2, Application US/10383930
; Publication No. US20040127400A1
; GENERAL INFORMATION:
; APPLICANT: Smith, Daniel J
; TITLE OF INVENTION: Immunogenicity of Glucan Binding Protein
; FILE REFERENCE: 25669-018
; CURRENT APPLICATION NUMBER: US/10/383,930
; CURRENT FILING DATE: 2003-03-07
; PRIOR APPLICATION NUMBER: 60/402,483
; PRIOR FILING DATE: 2002-08-08
; PRIOR APPLICATION NUMBER: 60/363,209
; PRIOR FILING DATE: 2002-03-07
; NUMBER OF SEQ ID NOS: 41
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 2
; LENGTH: 20
; TYPE: PRT
; ORGANISM: Streptococcus mutans
US-10-383-930-2

Query Match 100.0%; Score 97; DB 4; Length 20;
Best Local Similarity 100.0%; Pred. No. 3.6e-09;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KHLITIQGVSAALQTQQAG 20
|||
Db 1 KHLITIQGVSAALQTQQAG 20

RESULT 2

US-10-797-821-2
; Sequence 2, Application US/10797821
; Publication No. US20050031633A1
; GENERAL INFORMATION:
; APPLICANT: Smith, Daniel J.
; TITLE OF INVENTION: Glucan Binding Protein and Glycosyltransferase Immunogens
; FILE REFERENCE: 25669-020
; CURRENT APPLICATION NUMBER: US/10/797,821
; CURRENT FILING DATE: 2004-03-09
; PRIOR APPLICATION NUMBER: 10/383,930
; PRIOR FILING DATE: 2003-03-07
; PRIOR APPLICATION NUMBER: 60/363,209
; PRIOR FILING DATE: 2002-03-07
; PRIOR APPLICATION NUMBER: 60/402,483
; PRIOR FILING DATE: 2002-08-08
; PRIOR APPLICATION NUMBER: 09/290,049

; PRIOR FILING DATE: 1999-04-12
; PRIOR APPLICATION NUMBER: 60/081,550
; PRIOR FILING DATE: 1998-04-13
; PRIOR APPLICATION NUMBER: 60/115,142
; PRIOR FILING DATE: 1999-01-08
; NUMBER OF SEQ ID NOS: 45
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 2
; LENGTH: 20
; TYPE: PRT
; ORGANISM: Artificial
; FEATURE:
; OTHER INFORMATION: GpBB peptide
US-10-797-821-2

Query Match 100.0%; Score 97; DB 5; Length 20;
Best Local Similarity 100.0%; Pred. No. 3.6e-09;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KHKLTIQGVVSALQTQQAG 20
|||
DB 1 KHKLTIQGVVSALQTQQAG 20
|||

RESULT 3

US-10-383-930-9
; Sequence 9, Application US/10383930
; Publication No. US20040127400A1
; GENERAL INFORMATION:
; APPLICANT: Smith, Daniel J
; APPLICANT: Taubman, Martin A
; TITLE OF INVENTION: Immunogenicity of Glucan Binding Protein
; FILE REFERENCE: 25669-018
; CURRENT APPLICATION NUMBER: US/10/383,930
; CURRENT FILING DATE: 2003-03-07
; PRIOR APPLICATION NUMBER: 60/402,483
; PRIOR FILING DATE: 2002-08-08
; PRIOR APPLICATION NUMBER: 60/363,209
; PRIOR FILING DATE: 2002-03-07
; NUMBER OF SEQ ID NOS: 41
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 9
; LENGTH: 20
; TYPE: PRT
; ORGANISM: Streptococcus mutans
US-10-383-930-9

Query Match 67.0%; Score 65; DB 4; Length 20;
Best Local Similarity 100.0%; Pred. No. 0.00091;
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 6 TIQGVVSALQTQQA 19
|||
DB 6 TIQGVVSALQTQQA 19
|||

RESULT 4

US-10-797-821-9
; Sequence 9, Application US/10797821
; Publication No. US20050031633A1
; GENERAL INFORMATION:
; APPLICANT: Smith, Daniel J.
; APPLICANT: Taubman, Martin A.
; TITLE OF INVENTION: Glucan Binding Protein and Glycosyltransferase Immunogens
; FILE REFERENCE: 25669-020
; CURRENT APPLICATION NUMBER: US/10/797,821
; CURRENT FILING DATE: 2004-03-09
; PRIOR APPLICATION NUMBER: 10/383,930
; PRIOR FILING DATE: 2003-03-07
; PRIOR APPLICATION NUMBER: 60/363,209
; PRIOR FILING DATE: 2002-03-07
; PRIOR APPLICATION NUMBER: 60/402,483
; PRIOR FILING DATE: 2002-08-08

; PRIOR APPLICATION NUMBER: 09/290,049
; PRIOR FILING DATE: 1999-04-12
; PRIOR APPLICATION NUMBER: 60/081,550
; PRIOR FILING DATE: 1998-04-13
; PRIOR APPLICATION NUMBER: 60/115,142
; PRIOR FILING DATE: 1999-01-08
; NUMBER OF SEQ ID NOS: 45
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 9
; LENGTH: 20
; TYPE: PRT
; ORGANISM: Artificial
; FEATURE:
; OTHER INFORMATION: GpBB peptide
US-10-797-821-9

Query Match 67.0%; Score 65; DB 5; Length 20;
Best Local Similarity 100.0%; Pred. No. 0.00091;
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 6 TIQGVVSALQTQQA 19
|||
DB 6 TIQGVVSALQTQQA 19
|||

RESULT 5

US-10-383-930-29
; Sequence 29, Application US/10383930
; Publication No. US20040127400A1
; GENERAL INFORMATION:
; APPLICANT: Smith, Daniel J
; APPLICANT: Taubman, Martin A
; TITLE OF INVENTION: Immunogenicity of Glucan Binding Protein
; FILE REFERENCE: 25669-018
; CURRENT APPLICATION NUMBER: US/10/383,930
; CURRENT FILING DATE: 2003-03-07
; PRIOR APPLICATION NUMBER: 60/402,483
; PRIOR FILING DATE: 2002-08-08
; PRIOR APPLICATION NUMBER: 60/363,209
; PRIOR FILING DATE: 2002-03-07
; NUMBER OF SEQ ID NOS: 41
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 29
; LENGTH: 431
; TYPE: PRT
; ORGANISM: Streptococcus mutans
US-10-383-930-29

Query Match 67.0%; Score 65; DB 4; Length 431;
Best Local Similarity 100.0%; Pred. No. 0.0029; 0; Indels 0; Gaps 0;
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 6 TIQGVVSALQTQQA 19
|||
DB 57 TIQGVVSALQTQQA 70
|||

RESULT 6

US-10-383-930-30
; Sequence 30, Application US/10383930
; Publication No. US20040127400A1
; GENERAL INFORMATION:
; APPLICANT: Smith, Daniel J
; APPLICANT: Taubman, Martin A
; TITLE OF INVENTION: Immunogenicity of Glucan Binding Protein
; FILE REFERENCE: 25669-018
; CURRENT APPLICATION NUMBER: US/10/383,930
; CURRENT FILING DATE: 2003-03-07
; PRIOR APPLICATION NUMBER: 60/402,483
; PRIOR FILING DATE: 2002-08-08
; PRIOR APPLICATION NUMBER: 60/363,209
; PRIOR FILING DATE: 2002-03-07
; NUMBER OF SEQ ID NOS: 41

```

; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 30
; LENGTH: 431
; TYPE: PRT
; ORGANISM: Streptococcus mutans
US-10-383-930-30

```

Query Match 67.0%; Score 65; DB 4; Length 431;
Best Local Similarity 100.0%; Pred.No. 0.029;
Matches 14; Conservative 0; Mismatches 0; Indels

Qy 6 TIQQVSALQTQA 19
Db 57 TIQGVVSALQTQA 70

```

RESULT 7
US-10-383-930-33
; Sequence 33, Application US/10383930
; Publication No. US20040127400A1
; GENERAL INFORMATION:
; APPLICANT: Smith, Daniel J
; APPLICANT: Taubman, Martin A
; TITLE OF INVENTION: Immunogenicity of Glucan Binding Protein
; FILE REFERENCE: 25669-018
; CURRENT APPLICATION NUMBER: US/10/383,930
; CURRENT FILING DATE: 2003-03-07
; PRIOR APPLICATION NUMBER: 60/402,483
; PRIOR FILING DATE: 2002-08-08
; PRIOR APPLICATION NUMBER: 60/363,209
; PRIOR FILING DATE: 2002-03-07
; NUMBER OF SEQ ID NOS: 41
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 33
; LENGTH: 431
; TYPE: PRT
; ORGANISM: Streptococcus mutans
US-10-383-930-33

```

Query Match 67.0%; Score 65; DB 4; Length 431;
Best Local Similarity 100.0%; Pred.No. 0.029;
Matches 14; Conservative 0; Mismatches 0; Indels

Qy 6 TIQQVSA LQTQA 19
Db 57 TIQGVSA LQTQA 70

RESULT 8
US-10-797-821-29
; Sequence 29, Application US/10797821
; Publication No. US20050031633A1
; GENERAL INFORMATION:
; APPLICANT: Smith, Daniel J.
; APPLICANT: Taubman, Martin A.
; TITLE OF INVENTION: Glucan Binding Protein and Glycosyltransferase Immunogens
; FILE REFERENCE: 25669-020
; CURRENT APPLICATION NUMBER: US/10/797,821
; CURRENT FILING DATE: 2004-03-09
; PRIOR APPLICATION NUMBER: 10/383,930
; PRIOR FILING DATE: 2003-03-07
; PRIOR APPLICATION NUMBER: 60/363,209
; PRIOR FILING DATE: 2002-03-07
; PRIOR APPLICATION NUMBER: 60/402,483
; PRIOR FILING DATE: 2002-08-08
; PRIOR APPLICATION NUMBER: 09/290,049
; PRIOR FILING DATE: 1999-04-12
; PRIOR APPLICATION NUMBER: 60/081,550
; PRIOR FILING DATE: 1998-04-13
; PRIOR APPLICATION NUMBER: 60/115,142
; PRIOR FILING DATE: 1999-01-08
; NUMBER OF SEQ ID NOS: 45
; SOFTWARE: PatentIn version 3.2

```

; SEQ ID NO 29
; LENGTH: 431
; TYPE: prt
; ORGANISM: Streptococcus mutans
US-10-797-821-29

```

Query Match 67.0%; Score 65; DB 5; Length 431;
Best Local Similarity 100.0%; Pred.No. 0.029;
Matches 14; Conservative 0; Mismatches 0; Indels

Qy 6 TIQQVSALQTQQA 19
Db 57 TIQQVSALQTQQA 70

```

RESULT 9
US-10-797-821-30
; Sequence 30, Application US/10797821
; Publication No. US20050031633A1
; GENERAL INFORMATION:
; APPLICANT: Smith, Daniel J.
; APPLICANT: Taubman, Martin A.
; TITLE OF INVENTION: Glucan Binding Protein and Glycosyltransferase Immunogens
; FILE REFERENCE: 25669-020
; CURRENT APPLICATION NUMBER: US/10/797,821
; CURRENT FILING DATE: 2004-03-09
; PRIOR APPLICATION NUMBER: 10/383,930
; PRIOR FILING DATE: 2003-03-07
; PRIOR APPLICATION NUMBER: 60/363,209
; PRIOR FILING DATE: 2002-03-07
; PRIOR APPLICATION NUMBER: 60/402,483
; PRIOR FILING DATE: 2002-08-08
; PRIOR APPLICATION NUMBER: 09/290,049
; PRIOR FILING DATE: 1999-04-12
; PRIOR APPLICATION NUMBER: 60/081,550
; PRIOR FILING DATE: 1998-04-13
; PRIOR APPLICATION NUMBER: 60/115,142
; PRIOR FILING DATE: 1999-01-08
; NUMBER OF SEQ ID NOS: 45
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 30
; LENGTH: 431
; TYPE: PRT
; ORGANISM: Streptococcus mutans
US-10-797-821-30

```

Query Match 67.0%; Score 65; DB 5; Length 431;
Best Local Similarity 100.0%; Pred. No. 0.029;
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 6 TIQGVSAQTQA 19
|||
Db 57 TIQGVSAQTQA 70

RESULT 10
US-10-797-821-33
; Sequence 33, Application US/10797821
; Publication No. US20050031633A1
; GENERAL INFORMATION:
; APPLICANT: Smith, Daniel J.
; APPLICANT: Taubman, Martin A.
; TITLE OF INVENTION: Glucan Binding Protein and Glycosyltransferase Immunogens
; FILE REFERENCE: 25669-020
; CURRENT APPLICATION NUMBER: US/10/797, 821
; CURRENT FILING DATE: 2004-03-09
; PRIOR APPLICATION NUMBER: 10/383, 930
; PRIOR FILING DATE: 2003-03-07
; PRIOR APPLICATION NUMBER: 60/363, 209
; PRIOR FILING DATE: 2002-03-07
; PRIOR APPLICATION NUMBER: 60/402, 483
; PRIOR FILING DATE: 2002-08-08
; PRIOR APPLICATION NUMBER: 09/290, 049

```

; PRIOR FILING DATE: 1999-04-12
; PRIOR APPLICATION NUMBER: 60/081,550
; PRIOR FILING DATE: 1998-04-13
; PRIOR APPLICATION NUMBER: 60/115,142
; PRIOR FILING DATE: 1999-01-08
; NUMBER OF SEQ ID NOS: 45
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 33
; LENGTH: 431
; TYPE: PR1
; ORGANISM: Streptococcus mutans
US-10-797-821-33

```

```
Query Match      67.0%; Score 65; DB 5; Length 431;
Best Local Similarity 100.0%; Pred. No. 0.029;
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

Qy 6 TIQQVSA LQTQQA 19
Db 57 TIQGVSA LQTQQA 70

```

RESULT 11
US-10-383-930-31
; Sequence 31, Application US/10383930
; Publication No. US20040127400A1
; GENERAL INFORMATION:
; APPLICANT: Smith, Daniel J
; APPLICANT: Taubman, Martin A
; TITLE OF INVENTION: Immunogenicity of Glucan Binding Protein
; FILE REFERENCE: 25669-018
; CURRENT APPLICATION NUMBER: US/10/383,930
; CURRENT FILING DATE: 2003-03-07
; PRIOR APPLICATION NUMBER: 60/402,483
; PRIOR FILING DATE: 2002-08-08
; PRIOR APPLICATION NUMBER: 60/363,209
; PRIOR FILING DATE: 2002-03-07
; NUMBER OF SEQ ID NOS: 41
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 31
; LENGTH: 432
; TYPE: PRT
; ORGANISM: Streptococcus mutans
US-10-383-930-31

```

Query Match 67.0%; Score 65; DB 4; Length 432;
Best Local Similarity 100.0%; Pred. No. 0.029;
Matches 14: Conservative 0; Mismatches 0; Indels 0; Gaps 0

Qy 6 TIQGVSALQTQA 19
|||
Db 57 TIQGVSALOTQQA 70

```

RESULT 12
US-10-383-930-32
; Sequence 32, Application US/10383930
; Publication NO. US20040127400A1
; GENERAL INFORMATION:
; APPLICANT: Smith, Daniel J
; APPLICANT: Taubman, Martin A
; TITLE OF INVENTION: Immunogenicity of Glucan Binding Protein
; FILE REFERENCE: 25669-018
; CURRENT APPLICATION NUMBER: US/10/383,930
; CURRENT FILING DATE: 2003-03-07
; PRIOR APPLICATION NUMBER: 60/402,483
; PRIOR FILING DATE: 2002-08-08
; PRIOR APPLICATION NUMBER: 60/363,209
; PRIOR FILING DATE: 2002-03-07
; NUMBER OF SEQ ID NOS: 41
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 32
; LENGTH: 432

```

```

; TYPE: PRT
; ORGANISM: Streptococcus mutans
US-10-383-930-32

```

Query Match 67.0%; Score 65; DB 4; Length 432;
Best Local Similarity 100.0%; Pred. No. 0.029;
Matches 14; Conservative 0; Mismatches 0; Indels

Qy 6 TIQQVSA LQTQA 19
|||
Db 57 TIQGVSA LQTQA 70

```

RESULT 13
US-10-797-821-31
; Sequence 31, Application US/10797821
; Publication No. US20050031633A1
; GENERAL INFORMATION:
; APPLICANT: Smith, Daniel J.
; APPLICANT: Taubman, Martin A.
; TITLE OF INVENTION: Glucan Binding Protein and Glycosyltransferase Immunogens
; FILE REFERENCE: 25669-020
; CURRENT APPLICATION NUMBER: US/10/797,821
; CURRENT FILING DATE: 2004-03-09
; PRIOR APPLICATION NUMBER: 10/383,930
; PRIOR FILING DATE: 2003-03-07
; PRIOR APPLICATION NUMBER: 60/363,209
; PRIOR FILING DATE: 2002-03-07
; PRIOR APPLICATION NUMBER: 60/402,483
; PRIOR FILING DATE: 2002-08-08
; PRIOR APPLICATION NUMBER: 09/290,049
; PRIOR FILING DATE: 1999-04-12
; PRIOR APPLICATION NUMBER: 60/081,550
; PRIOR FILING DATE: 1998-04-13
; PRIOR APPLICATION NUMBER: 60/115,142
; PRIOR FILING DATE: 1999-01-08
; NUMBER OF SEQ ID NOS: 45
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 31
; LENGTH: 432
; TYPE: PRT
; ORGANISM: Streptococcus mutans
US-10-797-821-31

```

Query Match	67.0%;	Score 65;	DB 5;	Length 432;
Best Local Similarity	100.0%;	Pred. No. 0.029;		
Matches 14:	Conservative	0;	Mismatches	0;
	Indels	0;	Gaps	0;

Qy 6 TIQQVLSALQTQQA 19
|||
pb 57 TIOGOVSALOTQQA 70

RESULT 14

US-10-797-821-32

; Sequence 32, Application US/10797821

; Publication No. US20050031633A1

; GENERAL INFORMATION:

; APPLICANT: Smith, Daniel J.

; APPLICANT: Taubman, Martin A.

; TITLE OF INVENTION: Glucan Binding Protein and Glycosyltransferase Immunogens

; FILE REFERENCE: 25669-020

; CURRENT APPLICATION NUMBER: US/10/797,821

; CURRENT FILING DATE: 2004-03-09

; PRIOR APPLICATION NUMBER: 10/383,930

; PRIOR FILING DATE: 2003-03-07

; PRIOR APPLICATION NUMBER: 60/363,209

; PRIOR FILING DATE: 2002-03-07

; PRIOR APPLICATION NUMBER: 60/402,483

; PRIOR FILING DATE: 2002-08-08

; PRIOR APPLICATION NUMBER: 09/290,049

; PRIOR FILING DATE: 1999-04-12

; PRIOR APPLICATION NUMBER: 60/081,550

; PRIOR FILING DATE: 1998-04-13
 ; PRIOR APPLICATION NUMBER: 60/115,142
 ; PRIOR FILING DATE: 1999-01-08
 ; NUMBER OF SEQ ID NOS: 45
 ; SOFTWARE: PatentIn version 3.2
 ; SEQ ID NO 32
 ; LENGTH: 432
 ; TYPE: PRT
 ; ORGANISM: Streptococcus mutans
 ;
 US-10-797-821-32

Query Match 67.0%; Score 65; DB 5; Length 432;
 Best Local Similarity 100.0%; Pred. No. 0.029;
 Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 6 TIQGVVSALQTQQA 19
 |||||
 Db 57 TIQGVVSALQTQQA 70

RESULT 15
 US-10-383-930-8
 ; Sequence 8, Application US/10383930
 ; Publication No. US20040127400A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Smith, Daniel J
 ; APPLICANT: Taubman, Martin A
 ; TITLE OF INVENTION: Immunogenicity of Glucan Binding Protein
 ; FILE REFERENCE: 25669-018
 ; CURRENT APPLICATION NUMBER: US/10/383,930
 ; CURRENT FILING DATE: 2003-03-07
 ; PRIOR APPLICATION NUMBER: 60/402,483
 ; PRIOR FILING DATE: 2002-08-08
 ; PRIOR APPLICATION NUMBER: 60/363,209
 ; PRIOR FILING DATE: 2002-03-07
 ; NUMBER OF SEQ ID NOS: 41
 ; SOFTWARE: PatentIn version 3.2
 ; SEQ ID NO 8
 ; LENGTH: 20
 ; TYPE: PRT
 ; ORGANISM: Streptococcus mutans
 ;
 US-10-383-930-8

Query Match 52.6%; Score 51; DB 4; Length 20;
 Best Local Similarity 100.0%; Pred. No. 0.21;
 Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 6 TIQGVVSALQT 16
 |||||
 Db 10 TIQGVVSALQT 20

Search completed: January 25, 2006, 20:10:01
 Job time : 61 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2006 Compugen Ltd.

OM protein - protein search, using sw model

Run on: January 25, 2006, 19:15:26 ; Search time 5.5 Seconds
(without alignments)
39.378 Million cell updates/sec

Title: US-10-797-821-2

Perfect score: 97

Sequence: 1 KHKLTIQGVVSALQTQQA 20

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 75621 seqs, 10829074 residues

Total number of hits satisfying chosen parameters: 75621

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications AA New.*

- 1: /cgn2_6/prodata/2/pubpaa/US08_NEW_PUB.pep.*
- 2: /cgn2_6/prodata/2/pubpaa/US06_NEW_PUB.pep.*
- 3: /cgn2_6/prodata/2/pubpaa/US07_NEW_PUB.pep.*
- 4: /cgn2_6/prodata/2/pubpaa/PCT_NEW_PUB.pep.*
- 5: /cgn2_6/prodata/2/pubpaa/US09_NEW_PUB.pep.*
- 6: /cgn2_6/prodata/2/pubpaa/US10_NEW_PUB.pep.*
- 7: /cgn2_6/prodata/2/pubpaa/US11_NEW_PUB.pep.*
- 8: /cgn2_6/prodata/2/pubpaa/US60_NEW_PUB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	65	67.0	431	7	US-11-052-554A-210
2	40	41.2	398	7	US-11-052-554A-252
3	40	41.2	690	7	US-11-052-554A-232
4	39	40.2	1259	6	US-10-467-657-5510
5	38	39.2	318	7	US-11-194-246-294
6	38	39.2	1313	7	US-11-091-668-4
7	37	38.1	244	6	US-10-467-657-3418
8	37	38.1	804	6	US-10-467-962B-57
9	36	37.1	219	6	US-10-714-887-338
10	36	37.1	680	6	US-10-467-657-7612
11	36	37.1	716	6	US-10-467-657-8370
12	36	37.1	1141	6	US-10-995-561-1009
13	36	37.1	1141	6	US-10-995-561-1010
14	35	36.1	177	6	US-10-467-657-1282
15	35	36.1	202	6	US-10-454-437-62
16	35	36.1	475	7	US-11-074-176-336
17	35	36.1	488	7	US-11-074-176-126
18	35	36.1	529	6	US-10-858-730-104
19	35	36.1	529	6	US-10-858-730-105
20	35	36.1	642	6	US-10-467-657-5304
21	35	36.1	781	7	US-11-194-246-344
22	34.5	35.6	5405	7	US-11-108-172-1116
23	34	35.1	76	6	US-10-510-880-6
24	34	35.1	76	6	US-10-977-334-2
25	34	35.1	108	6	US-10-510-880-5

26	34	35.1	108	6	US-10-875-800-1	Sequence 1, Appli
27	34	35.1	108	6	US-10-299-977-1	Sequence 1, Appli
28	34	35.1	134	6	US-10-510-880-1	Sequence 1, Appli
29	34	35.1	134	6	US-10-875-800-2	Sequence 2, Appli
30	34	35.1	134	7	US-11-043-590-12	Sequence 12, Appli
31	34	35.1	161	6	US-10-467-657-22	Sequence 22, Appli
32	34	35.1	161	6	US-10-467-657-6198	Sequence 6198, Ap
33	34	35.1	162	7	US-11-043-590-13	Sequence 13, Appli
34	34	35.1	206	6	US-10-467-657-7716	Sequence 7716, Ap
35	34	35.1	307	7	US-11-055-822-528	Sequence 528, App
36	34	35.1	336	6	US-10-821-234-957	Sequence 957, App
37	34	35.1	373	6	US-10-793-626-2288	Sequence 2288, Ap
38	34	35.1	665	6	US-10-793-626-1232	Sequence 1232, Ap
39	34	35.1	1378	7	US-11-212-443-60	Sequence 60, Appl
40	33.5	34.5	3674	7	US-11-000-463-454	Sequence 454, App
41	33	34.0	19	6	US-10-503-575-254	Sequence 254, App
42	33	34.0	156	6	US-10-508-263-106	Sequence 106, App
43	33	34.0	172	7	US-11-156-084-140	Sequence 140, App
44	33	34.0	217	7	US-11-056-408-12	Sequence 12, Appl
45	33	34.0	266	7	US-11-000-463-464	Sequence 464, App

ALIGNMENTS

RESULT 1
US-11-052-554A-210
; Sequence 210, Application US/11052554A
; Publication No. US20050288866A1
; GENERAL INFORMATION:
; APPLICANT: Sachdeva, et al.
; TITLE OF INVENTION: COMPUTATIONAL METHOD FOR IDENTIFYING ADHESIN AND ADHESIN-LIKE
; TITLE OF INVENTION: PROTEINS OF THERAPEUTIC POTENTIAL
; FILE REFERENCE: 30853/40359A
; CURRENT APPLICATION NUMBER: US/11/052,554A
; CURRENT FILING DATE: 2005-02-07
; PRIOR APPLICATION NUMBER: US 60/589,227
; PRIOR FILING DATE: 2004-07-20
; PRIOR APPLICATION NUMBER: IN 173/DEL/2004
; PRIOR FILING DATE: 2004-02-06
; NUMBER OF SEQ ID NOS: 763
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 210
; LENGTH: 431
; TYPE: PRT
; ORGANISM: Streptococcus mutans UA159
US-11-052-554A-210

Query Match 67.0%; Score 65; DB 7; Length 431;
Best Local Similarity 100.0%; Pred. No. 0.00048;
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy	6	TIQGVVSALQTQQA	19
Db	57	TIQGVVSALQTQQA	70

RESULT 2
US-11-052-554A-252
; Sequence 252, Application US/11052554A
; Publication No. US20050288866A1
; GENERAL INFORMATION:
; APPLICANT: Sachdeva, et al.
; TITLE OF INVENTION: COMPUTATIONAL METHOD FOR IDENTIFYING ADHESIN AND ADHESIN-LIKE
; TITLE OF INVENTION: PROTEINS OF THERAPEUTIC POTENTIAL
; FILE REFERENCE: 30853/40359A
; CURRENT APPLICATION NUMBER: US/11/052,554A
; CURRENT FILING DATE: 2005-02-07
; PRIOR APPLICATION NUMBER: US 60/589,227
; PRIOR FILING DATE: 2004-07-20
; PRIOR APPLICATION NUMBER: IN 173/DEL/2004
; PRIOR FILING DATE: 2004-02-06
; NUMBER OF SEQ ID NOS: 763

; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 252
; LENGTH: 398
; TYPE: PRT
; ORGANISM: Streptococcus pyogenes MGAS8232
US-11-052-554A-252

Query Match 41.2%; Score 40; DB 7; Length 398;
Best Local Similarity 38.9%; Pred. No. 12;
Matches 7; Conservative 8; Mismatches 3; Indels 0; Gaps 0;

QY 1 KHKLITIQGVSAQTQQ 18
DB 49 QNQSALQAQVSSLSQ 66

RESULT 3

US-11-052-554A-232
; Sequence 232, Application US/11052554A
; Publication No. US20050288866A1
; GENERAL INFORMATION:
; APPLICANT: Sachdeva, et al.
; TITLE OF INVENTION: COMPUTATIONAL METHOD FOR IDENTIFYING ADHESIN AND ADHESIN-LIKE
; TITLE OF INVENTION: PROTEINS OF THERAPEUTIC POTENTIAL
; FILE REFERENCE: 30853/40359A
; CURRENT APPLICATION NUMBER: US/11/052,554A
; PRIOR FILING DATE: 2005-02-07
; PRIOR APPLICATION NUMBER: US 60/589,227
; PRIOR FILING DATE: 2004-07-20
; PRIOR APPLICATION NUMBER: IN 173/DEL/2004
; PRIOR FILING DATE: 2004-02-06
; NUMBER OF SEQ ID NOS: 763
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 232
; LENGTH: 690
; TYPE: PRT
; ORGANISM: Streptococcus pneumoniae R6
US-11-052-554A-232

Query Match 41.2%; Score 40; DB 7; Length 690;
Best Local Similarity 33.3%; Pred. No. 23;
Matches 5; Conservative 7; Mismatches 3; Indels 0; Gaps 0;

QY 1 KHKLITIQGVSAQ 15
DB 60 KYKVLIVGEGNIGTVQ 74

RESULT 4

US-10-467-657-5510
; Sequence 5510, Application US/10467657
; Publication No. US20050260581A1
; GENERAL INFORMATION:
; APPLICANT: CHIRON SpA
; APPLICANT: FONTANA Maria Rita
; APPLICANT: PIZZA Mariagrazia
; APPLICANT: MASIGNANI Vega
; APPLICANT: MONACI Elisabetta
; TITLE OF INVENTION: GONOCOCCAL PROTEINS AND NUCLEIC ACIDS
; FILE REFERENCE:
; CURRENT APPLICATION NUMBER: US/10/467,657
; CURRENT FILING DATE: 2003-08-11
; PRIOR FILING DATE: GB-0103424.8
; PRIOR FILING DATE: 2001-02-12
; NUMBER OF SEQ ID NOS: 9218
; SOFTWARE: SeqWin99, version 1.04
; SEQ ID NO 5510
; LENGTH: 1259
; TYPE: PRT
; ORGANISM: Neisseria gonorrhoeae
US-10-467-657-5510

Query Match 40.2%; Score 39; DB 6; Length 1259;

Best Local Similarity 47.1%; Pred. No. 71;
Matches 8; Conservative 3; Mismatches 6; Indels 0; Gaps 0;
QY 2 KKLITIQGVSAQTQQ 18
DB 349 HKPIMIAGGLGSIQAQQ 365

RESULT 5

US-11-194-246-294
; Sequence 294, Application US/11194246
; Publication No. US20050272089A1
; GENERAL INFORMATION:
; APPLICANT: Mott, John
; APPLICANT: Trepod, Catherine
; APPLICANT: Arvidson, Staffan
; TITLE OF INVENTION: CRITICAL GENES AND POLYPEPTIDES OF HAEMOPHILUS INFLUENZAE AND MET
; TITLE OF INVENTION: USE
; FILE REFERENCE: 00592.US1 (M&R 268.05920101)
; CURRENT APPLICATION NUMBER: US/11/194,246
; CURRENT FILING DATE: 2005-08-01
; PRIOR FILING DATE: 2002-10-21
; PRIOR APPLICATION NUMBER: US 60/345,438
; PRIOR FILING DATE: 2001-10-19
; NUMBER OF SEQ ID NOS: 621
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 294
; LENGTH: 318
; TYPE: PRT
; ORGANISM: HAEMOPHILUS INFLUENZAE
US-11-194-246-294

Query Match 39.2%; Score 38; DB 7; Length 318;
Best Local Similarity 38.9%; Pred. No. 20;
Matches 7; Conservative 5; Mismatches 6; Indels 0; Gaps 0;

QY 3 KLITIQGVSAQTQQ 20
DB 277 QIATVDGVNLGLQSQ 294

RESULT 6

US-11-091-668-4
; Sequence 4, Application US/11091668
; Publication No. US20050262585A1
; GENERAL INFORMATION:
; APPLICANT: University of Nebraska
; APPLICANT: Mackenzie, Sally Ann
; APPLICANT: Vaghchhipawala, Zarir Erach
; TITLE OF INVENTION: Soybean FGAM Synthase Promoters Useful In Parasite Control
; FILE REFERENCE: 1231-221
; CURRENT APPLICATION NUMBER: US/11/091,668
; CURRENT FILING DATE: 2005-03-28
; PRIOR APPLICATION NUMBER: 60556745
; PRIOR FILING DATE: 2004-03-26
; NUMBER OF SEQ ID NOS: 18
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 4
; LENGTH: 1313
; TYPE: PRT
; ORGANISM: Glycine max
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (806)..(806)
; OTHER INFORMATION: The 'Xaa' at location 806 stands for Gly.
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (812)..(812)
; OTHER INFORMATION: The 'Xaa' at location 812 stands for Val.
US-11-091-668-4

Query Match 39.2%; Score 38; DB 7; Length 1313;

Best Local Similarity 38.9%; Pred. No. 1.1e+02; Mismatches 4; Indels 0; Gaps 0;

QY 1 KHLITIQGVSAALQTOQ 18
 Db 11 KEAQVKISQIVIEIQTEQ 28

RESULT 7

US-10-467-657-3418
 ; Sequence 3418, Application US/10467657
 ; Publication No. US20050260581A1
 ; GENERAL INFORMATION:
 ; APPLICANT: CHIRON SpA
 ; APPLICANT: FONTANA Maria Rita
 ; APPLICANT: PIZZA Mariagrazia
 ; APPLICANT: MASIGNANI Vega
 ; APPLICANT: MONACI Elisabetta
 ; TITLE OF INVENTION: GONOCOCCAL PROTEINS AND NUCLEIC ACIDS
 ; FILE REFERENCE:
 ; CURRENT APPLICATION NUMBER: US/10/467,657
 ; CURRENT FILING DATE: 2003-08-11
 ; PRIOR APPLICATION NUMBER: GB-0103424.8
 ; PRIOR FILING DATE: 2001-02-12
 ; NUMBER OF SEQ ID NOS: 9218
 ; SOFTWARE: Seqwin99, version 1.04
 ; SEQ ID NO 3418
 ; LENGTH: 244
 ; TYPE: PRT
 ; ORGANISM: Neisseria gonorrhoeae
 US-10-467-657-3418

Query Match 38.1%; Score 37; DB 6; Length 244;
 Best Local Similarity 42.1%; Pred. No. 22;
 Matches 8; Conservative 5; Mismatches 6; Indels 0; Gaps 0;

QY 2 HKLITIQGVSAALQTOQAG 20
 Db 67 NRVFTLLGGTALQKGQAG 85

RESULT 8

US-10-467-962B-57
 ; Sequence 57, Application US/10467962B
 ; Publication No. US20050246784A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Plesch, Gunnar
 ; APPLICANT: Blau, Astrid
 ; APPLICANT: Daeschner, Klaus
 ; APPLICANT: Klein, Mathieu
 ; TITLE OF INVENTION: Identification of Herbicidally Active Substances
 ; FILE REFERENCE: 2000 857
 ; CURRENT APPLICATION NUMBER: US/10/467,962B
 ; CURRENT FILING DATE: 2003-08-14
 ; PRIOR APPLICATION NUMBER: PCT/EP02/01466
 ; PRIOR FILING DATE: 2002-02-13
 ; NUMBER OF SEQ ID NOS: 109
 ; SOFTWARE: PatentIn Vers. 2.0
 ; SEQ ID NO 57
 ; LENGTH: 804
 ; TYPE: PRT
 ; ORGANISM: Arabidopsis thaliana
 US-10-467-962B-57

Query Match 38.1%; Score 37; DB 6; Length 804;
 Best Local Similarity 60.0%; Pred. No. 93;
 Matches 9; Conservative 1; Mismatches 5; Indels 0; Gaps 0;

QY 6 TIQGVSAALQTOQAG 20
 Db 616 TSQGVSAAGKLEAG 630

RESULT 9

US-10-714-887-338
 ; Sequence 338, Application US/10714887
 ; Publication No. US20060015972A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Mendel Biotechnology, Inc.
 ; APPLICANT: HEARD, Jacqueline
 ; APPLICANT: RIECHMANN, Jose Luis
 ; APPLICANT: CREELMAN, Robert
 ; APPLICANT: RATCLIFFE, Oliver
 ; APPLICANT: CANALES, Roger
 ; APPLICANT: REPETTI, Peter
 ; APPLICANT: KUMIMOTO, Roderick W
 ; APPLICANT: GUTTERSON, Neal
 ; APPLICANT: REUBER, T. Lynne
 ; APPLICANT: PINEDA, Omaira
 ; APPLICANT: SHERMAN, Bradley K
 ; TITLE OF INVENTION: PLANT TRANSCRIPTIONAL REGULATORS OF DROUGHT STRESS
 ; FILE REFERENCE: MB10058-CIP
 ; CURRENT APPLICATION NUMBER: US/10/714,887
 ; CURRENT FILING DATE: 2003-11-13
 ; PRIOR APPLICATION NUMBER: 10/412,699
 ; PRIOR FILING DATE: 2003-04-10
 ; PRIOR APPLICATION NUMBER: 09/506,720
 ; PRIOR FILING DATE: 2000-02-17
 ; PRIOR APPLICATION NUMBER: 60/135,134
 ; PRIOR FILING DATE: 1999-05-20
 ; PRIOR APPLICATION NUMBER: 09/394,519
 ; PRIOR FILING DATE: 1999-09-13
 ; PRIOR APPLICATION NUMBER: 09/533,392
 ; PRIOR FILING DATE: 2000-03-22
 ; PRIOR APPLICATION NUMBER: 09/533,029
 ; PRIOR FILING DATE: 2000-03-22
 ; PRIOR APPLICATION NUMBER: 09/532,591
 ; PRIOR FILING DATE: 2000-03-22
 ; PRIOR APPLICATION NUMBER: 09/533,030
 ; PRIOR FILING DATE: 2000-03-22
 ; PRIOR APPLICATION NUMBER: 60/125,814
 ; PRIOR FILING DATE: 1999-03-23
 ; PRIOR APPLICATION NUMBER: 09/713,994
 ; PRIOR FILING DATE: 2000-11-16
 ; Remaining Prior Application data removed - See File Wrapper or PALM.
 ; NUMBER OF SEQ ID NOS: 430
 ; SOFTWARE: PatentIn version 3.2
 ; SEQ ID NO 338
 ; LENGTH: 219
 ; TYPE: PRT
 ; ORGANISM: Oryza sativa (japonica cultivar-group)
 ; FEATURE:
 ; OTHER INFORMATION: G3397 polypeptide Orthologous to G481 and G482
 US-10-714-887-338

Query Match 37.1%; Score 36; DB 6; Length 219;
 Best Local Similarity 42.1%; Pred. No. 28;
 Matches 8; Conservative 3; Mismatches 8; Indels 0; Gaps 0;

QY 2 HKLITIQGVSAALQTOQAG 20
 Db 109 HKFREIEGERAAATTGAG 127

RESULT 10

US-10-467-657-7612
 ; Sequence 7612, Application US/10467657
 ; Publication No. US20050260581A1
 ; GENERAL INFORMATION:
 ; APPLICANT: CHIRON SpA
 ; APPLICANT: FONTANA Maria Rita
 ; APPLICANT: PIZZA Mariagrazia
 ; APPLICANT: MASIGNANI Vega
 ; APPLICANT: MONACI Elisabetta
 ; TITLE OF INVENTION: GONOCOCCAL PROTEINS AND NUCLEIC ACIDS
 ; FILE REFERENCE:

```
; CURRENT APPLICATION NUMBER: US/10/467,657
; CURRENT FILING DATE: 2003-08-11
; PRIOR APPLICATION NUMBER: GB-0103424.8
; PRIOR FILING DATE: 2001-02-12
; NUMBER OF SEQ ID NOS: 9218
; SOFTWARE: SeqWin99, version 1.04
; SEQ ID NO 7612
; LENGTH: 680
; TYPE: PRT
; ORGANISM: Neisseria gonorrhoeae
US-10-467-657-7612

Query Match          37.1%; Score 36; DB 6; Length 680;
Best Local Similarity 35.3%; Pred. No. 1.1e+02;
Matches 6; Conservative 7; Mismatches 4; Indels 0; Gaps 0;

QY      4 LITIQGVSAALQTQAG 20
Db      628 LDTLRGELGTLTRSSG 644

RESULT 11
US-10-467-657-8370
; Sequence 8370, Application US/10467657
; Publication No. US20050260581A1
; GENERAL INFORMATION:
; APPLICANT: CHIRON SpA
; APPLICANT: FONTANA Maria Rita
; APPLICANT: PIZZA Mariagrazia
; APPLICANT: MASIGNANI Vega
; APPLICANT: MONACI Elisabetta
; TITLE OF INVENTION: GONOCOCCAL PROTEINS AND NUCLEIC ACIDS
; FILE REFERENCE:
; CURRENT APPLICATION NUMBER: US/10/467,657
; CURRENT FILING DATE: 2003-08-11
; PRIOR APPLICATION NUMBER: GB-0103424.8
; PRIOR FILING DATE: 2001-02-12
; NUMBER OF SEQ ID NOS: 9218
; SOFTWARE: SeqWin99, version 1.04
; SEQ ID NO 8370
; LENGTH: 716
; TYPE: PRT
; ORGANISM: Neisseria gonorrhoeae
US-10-467-657-8370

Query Match          37.1%; Score 36; DB 6; Length 716;
Best Local Similarity 35.3%; Pred. No. 1.2e+02;
Matches 6; Conservative 7; Mismatches 4; Indels 0; Gaps 0;

QY      4 LITIQGVSAALQTQAG 20
Db      664 LDTLRGELGTLTRSSG 680

RESULT 12
US-10-995-561-1009
; Sequence 1009, Application US/10995561
; Publication No. US20050272054A1
; GENERAL INFORMATION:
; APPLICANT: CARGILL, Michele et al.
; TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
; TITLE OF INVENTION: CARDIOVASCULAR DISORDERS AND DRUG RESPONSE, METHODS OF
; FILE REFERENCE: CL001559
; CURRENT APPLICATION NUMBER: US/10/995,561
; CURRENT FILING DATE: 2004-11-24
; NUMBER OF SEQ ID NOS: 85702
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1009
; LENGTH: 1141
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-995-561-1009

Query Match          37.1%; Score 36; DB 6; Length 1141;
Best Local Similarity 50.0%; Pred. No. 2.1e+02;
Matches 7; Conservative 3; Mismatches 4; Indels 0; Gaps 0;

QY      5 ITIQGVSAALQTQ 18
Db      197 VTIIQQVQTVQAQR 210

RESULT 13
US-10-995-561-1010
; Sequence 1010, Application US/10995561
; Publication No. US20050272054A1
; GENERAL INFORMATION:
; APPLICANT: CARGILL, Michele et al.
; TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
; TITLE OF INVENTION: CARDIOVASCULAR DISORDERS AND DRUG RESPONSE, METHODS OF
; FILE REFERENCE: CL001559
; CURRENT APPLICATION NUMBER: US/10/995,561
; CURRENT FILING DATE: 2004-11-24
; NUMBER OF SEQ ID NOS: 85702
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1010
; LENGTH: 1141
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-995-561-1010

Query Match          37.1%; Score 36; DB 6; Length 1141;
Best Local Similarity 50.0%; Pred. No. 2.1e+02;
Matches 7; Conservative 3; Mismatches 4; Indels 0; Gaps 0;

QY      5 ITIQGVSAALQTQ 18
Db      197 VTIIQQVQTVQAQR 210

RESULT 14
US-10-467-657-1282
; Sequence 1282, Application US/10467657
; Publication No. US20050260581A1
; GENERAL INFORMATION:
; APPLICANT: CHIRON SpA
; APPLICANT: FONTANA Maria Rita
; APPLICANT: PIZZA Mariagrazia
; APPLICANT: MASIGNANI Vega
; APPLICANT: MONACI Elisabetta
; TITLE OF INVENTION: GONOCOCCAL PROTEINS AND NUCLEIC ACIDS
; FILE REFERENCE:
; CURRENT APPLICATION NUMBER: US/10/467,657
; CURRENT FILING DATE: 2003-08-11
; PRIOR APPLICATION NUMBER: GB-0103424.8
; PRIOR FILING DATE: 2001-02-12
; NUMBER OF SEQ ID NOS: 9218
; SOFTWARE: SeqWin99, version 1.04
; SEQ ID NO 1282
; LENGTH: 177
; TYPE: PRT
; ORGANISM: Neisseria gonorrhoeae
US-10-467-657-1282

Query Match          36.1%; Score 35; DB 6; Length 177;
Best Local Similarity 50.0%; Pred. No. 33;
Matches 6; Conservative 5; Mismatches 1; Indels 0; Gaps 0;

QY      5 ITIQGVSAALQT 16
Db      162 LSVQGLKLSALYT 173

RESULT 15
```

```

US-10-454-437-62
; Sequence 62, Application US/10454437
; Publication No. US2005027115A1
; GENERAL INFORMATION:
; APPLICANT: Pompejus, Markus
; APPLICANT: Kroger, Burkhard
; APPLICANT: Schroder, Hartwig
; APPLICANT: Zelder, Oskar
; APPLICANT: Habethauer, Gregor
; TITLE OF INVENTION: CORYNEBACTERIUM GLUTAMICUM GENES ENCODING PROTEINS
; TITLE OF INVENTION: INVOLVED IN HOMEOSTASIS AND ADAPTATION
; FILE REFERENCE: BGI-128CPCN
; CURRENT APPLICATION NUMBER: US/10/454,437
; CURRENT FILING DATE: 2003-06-13
; PRIOR APPLICATION NUMBER: US 60/141031
; PRIOR FILING DATE: 1999-06-25
; PRIOR APPLICATION NUMBER: DE 19931636.8
; PRIOR FILING DATE: 1999-07-08
; PRIOR APPLICATION NUMBER: DE 19932125.6
; PRIOR FILING DATE: 1999-07-09
; PRIOR APPLICATION NUMBER: DE 19932126.4
; PRIOR FILING DATE: 1999-07-09
; PRIOR APPLICATION NUMBER: DE 19932127.2
; PRIOR FILING DATE: 1999-07-09
; PRIOR APPLICATION NUMBER: DE 19932128.0
; PRIOR FILING DATE: 1999-07-09
; PRIOR APPLICATION NUMBER: DE 19932129.9
; PRIOR FILING DATE: 1999-07-19
; PRIOR APPLICATION NUMBER: DE 19932226.0
; PRIOR FILING DATE: 1999-07-09
; PRIOR APPLICATION NUMBER: DE 19932920.6
; PRIOR FILING DATE: 1999-07-14
; PRIOR APPLICATION NUMBER: DE 19932922.2
; PRIOR FILING DATE: 1999-07-14
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 442
; SEQ ID NO 62
; LENGTH: 202
; TYPE: PRT
; ORGANISM: Corynebacterium glutamicum
US-10-454-437-62

```

```

Query Match      36.1%; Score 35; DB 6; Length 202;
Best Local Similarity 66.7%; Pred. No. 39;
Matches 8; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

```

```

Qy      8 QGQVSALQTQQA 19
      |||||:|
Db     135 QGQVSDLEIQAA 146

```

Search completed: January 25, 2006, 20:11:07
Job time : 5.5 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2006 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: January 25, 2006, 18:48:24 : Search time 68.2 Seconds
(without alignments)
128.850 Million cell updates/sec

Title: US-10-797-821-3

Perfect score: 90

Sequence: 1 TATEAQPSSASTAAVAAN 20

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 2443163 seqs, 439378781 residues

Total number of hits satisfying chosen parameters: 2443163

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

1: Geneseq 21.*

2: Geneseqp1980s.*

3: Geneseqp1990s.*

4: Geneseqp2000s.*

5: Geneseqp2001s.*

6: Geneseqp2002s.*

7: Geneseqp2003as.*

8: Geneseqp2003bs.*

9: Geneseqp2004s.*

10: Geneseqp2005s.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	90	100.0	20	7	ADD93623 Streptoco
2	90	100.0	20	9	AD37246 Streptoco
3	90	100.0	431	7	ADD93649 Streptoco
4	90	100.0	431	7	ADD93650 Streptoco
5	90	100.0	431	9	AD37272 Streptoco
6	90	100.0	431	9	AD37273 Streptoco
7	90	100.0	431	9	AEB91500 Microbial
8	84	93.3	432	7	ADD93651 Streptoco
9	84	93.3	432	7	ADD93652 Streptoco
10	84	93.3	432	9	AD37274 Streptoco
11	84	93.3	432	9	AD37275 Streptoco
12	82	91.1	431	7	ADD93653 Streptoco
13	82	91.1	431	9	AD37276 Streptoco
14	67	74.4	20	9	ADD93638 Streptoco
15	67	74.4	20	9	AD37261 Streptoco
16	55	61.1	291	4	AB67831 Drosophil
17	54	60.0	1106	6	ABU17126 Protein e
18	49.5	55.0	371	8	AD391331 Plant ful
19	49	54.4	189	8	AD391493 Plant ful
20	49	54.4	231	8	ADY23867 Plant ful
21	48	53.3	132	3	AA661553 Arabidops
22	48	53.3	151	3	AA661552 Arabidops
23	48	53.3	151	3	AA661551 Arabidops
24	48	53.3	154	3	AA661550 Arabidops

ALIGNMENTS

RESULT 1

ADD93623

ID ADD93623 standard; peptide; 20 AA.

AC ADD93623;

XX

DT 29-JAN-2004 (first entry)

XX

DE Streptococcus mutans glucan binding protein-B SAS peptide.

XX

KW Glucan binding protein-B; GbpB; vaccine; anticaries; epitope; immunogen.

XX

OS Streptococcus mutans.

XX

PN WO2003075845-A2.

XX

PD 18-SEP-2003.

XX

PF 07-MAR-2003; 2003WO-US006962.

XX

PR 07-MAR-2002; 2002US-0363209P.

XX

PR 08-AUG-2002; 2002US-0402483P.

XX

XX (FORS-) FORSYTH INST.

PA Smith DJ, Taubman MA;

PI WPI; 2003-845091/78.

DR

XX Composition useful as vaccines for dental caries comprises a fragment of a glucan binding protein-B binding to a major histocompatibility complex class II protein.

PT

XX Claim 12; Page 10; 49pp; English.

PS

XX

CC The present sequence is that of SAS peptide comprising amino acid residues 306-325 of the glucan binding protein-B (GbpB) of Streptococcus mutans strain SK32 ADD93649. The peptide binds to a major histocompatibility complex (MHC) class II protein. It was identified as a potential B cell epitope using a matrix-based algorithm for epitope prediction, which was used to search the primary amino acid sequence of GbpB for known MHC class II binding motifs. The peptide can be used in immunogenic compositions and subunit vaccines for dental caries. These comprise MHC class II protein-binding GbpB peptides covalently linked with peptide subunits (preferably from the catalytic domain) of a glucosyltransferase. The compositions are used in a claimed method of eliciting production of an antibody in a mammal. Dieptopic or

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC multiepitopic polypeptides can be prepared synthetically or by
 CC recombinant DNA technology. Antibodies raised against MHC class II
 CC binding fragments of GbpB can be used in passive immunisation.

SQ Sequence 20 AA;
 Query Match 100.0%; Score 90; DB 7; Length 20;
 Best Local Similarity 100.0%; Pred. No. 2.1e-06;
 Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 TATEAQPSSASSASTAAVAAN 20
 Db 1 TATEAQPSSASSASTAAVAAN 20

RESULT 2
 ADX37246
 ID ADX37246 standard; peptide; 20 AA.
 XX AC ADX37246;
 XX DT 21-APR-2005 (first entry)
 XX DE Streptococcus mutant glucan binding protein B peptide #3.
 XX KW immunogenicity; immune stimulation; glucan binding protein-B;
 KW microparticle; major histocompatibility complex; tooth disease.
 XX OS Streptococcus mutans.
 XX PN US2005031633-A1.
 XX PD 10-FEB-2005.
 XX PF 09-MAR-2004; 2004US-00797821.
 XX PR 13-APR-1998; 98US-0081550P.
 PR 08-JAN-1999; 99US-0115142P.
 PR 12-APR-1999; 99US-00290049.
 PR 07-MAR-2002; 2002US-0363209P.
 PR 08-AUG-2002; 2002US-0402483P.
 PR 07-MAR-2003; 2003US-00383930.
 XX {SMIT/} SMITH D J.
 PA {TAUB/} TAUBMAN M A.
 XX Smith DJ, Taubman MA;
 PI WPI; 2005-151644/16.

XX New composition comprising a fragment of a glucan binding protein-B
 (GbpB) that binds to MHC class II protein, and a biocompatible
 PT microparticle, useful for producing an antibody (claimed) for immunizing
 PT mammals against dental caries.
 XX Claim 4; SEQ ID NO 3; 73pp; English.
 PS The invention relates to a composition comprising a fragment of a glucan
 CC binding protein-B (GbpB) and a biocompatible microparticle, where the
 CC fragment binds to a major histocompatibility complex (MHC) class II
 CC protein. The composition is useful for producing an antibody for
 CC immunizing mammals against dental caries. This sequence corresponds to a
 CC Streptococcus mutans GbpB-derived peptide of the invention.

SQ Sequence 20 AA;
 Query Match 100.0%; Score 90; DB 9; Length 20;
 Best Local Similarity 100.0%; Pred. No. 2.1e-06;
 Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 TATEAQPSSASSASTAAVAAN 20
 Db 1 TATEAQPSSASSASTAAVAAN 20

RESULT 3
 ADD93649
 ID ADD93649 standard; protein; 431 AA.
 XX AC ADD93649;
 XX DT 29-JAN-2004 (first entry)
 XX DE Streptococcus mutans glucan binding protein-B.
 KW Glucan binding protein-B; GbpB; vaccine; anticaries; epitope; immunogen.
 XX OS Streptococcus mutans.
 XX Key Location/Qualifiers
 FH Region 6..25
 FT /note= "HLA-binding peptide"
 FT Region 16..35
 FT /note= "HLA-binding peptide"
 FT Region 33..52
 FT /note= "HLA-binding peptide"
 FT Region 37..56
 FT /note= "HLA-binding peptide"
 FT Region 48..67
 FT /note= "HLA-binding peptide"
 FT Region 52..71
 FT /note= "HLA-binding peptide"
 FT Region 88..107
 FT /note= "HLA-binding peptide"
 FT Region 113..132
 FT /note= "HLA-binding peptide"
 FT Region 117..136
 FT /note= "HLA-binding peptide"
 FT Region 137..156
 FT /note= "HLA-binding peptide"
 FT Region 174..193
 FT /note= "HLA-binding peptide"
 FT Region 194..213
 FT /note= "HLA-binding peptide"
 FT Region 214..233
 FT /note= "HLA-binding peptide"
 FT Region 248..267
 FT /note= "HLA-binding peptide"
 FT Region 289..308
 FT /note= "HLA-binding peptide"
 FT Region 306..325
 FT /note= "HLA-binding peptide"
 FT Region 311..330
 FT /note= "HLA-binding peptide"
 FT Region 349..368
 FT /note= "HLA-binding peptide"
 FT Region 365..384
 FT /note= "HLA-binding peptide"
 FT Region 383..402
 FT /note= "HLA-binding peptide"
 FT Region 403..422
 FT /note= "HLA-binding peptide"
 XX WO2003075845-A2.
 XX 19-SEP-2003.
 XX 07-MAR-2003; 2003WO-US006962.
 XX 07-MAR-2002; 2002US-0363209P.
 PR 08-AUG-2002; 2002US-0402483P.
 XX (FORS-) FORSYTH INST.
 XX Smith DJ, Taubman MA;
 XX PI

DR WPI: 2003-845091/78.
 DR GENBANK; AY046410.
 XX
 CC Composition useful as vaccines for dental caries comprises a fragment of
 PT a glucan binding protein-B binding to a major histocompatibility complex
 PT class II protein.
 XX
 PS Claim 6; Page 7; 49pp; English.
 XX
 CC The present sequence is the protein sequence of the glucan binding
 CC protein-B (GbpB) of Streptococcus mutans strain SK32. The sequence
 CC includes a number of human leucocyte antigen (HLA)-binding epitopes. The
 CC invention provides immunogenic compositions and vaccines for dental
 CC caries. The compositions comprise major histocompatibility complex (MHC)
 CC class II protein-binding GbpB peptides, especially HLA-binding peptides,
 CC covalently linked with peptide subunits of a glucosyltransferase. The
 CC compositions are used in a claimed method of eliciting production of an
 CC antibody in a mammal. Diepitopic or multiepitopic polypeptides can be
 CC prepared synthetically or by recombinant DNA technology. Antibodies
 CC raised against MHC class II binding fragments of GbpB can be used in
 CC passive immunisation.
 XX
 SQ Sequence 431 AA;
 Query Match 100.0%; Score 90; DB 7; Length 431;
 Best Local Similarity 100.0%; Pred. No. 5.7e-05;
 Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 TATEAQPSSASSASTAAVAAN 20
 Db 306 TATEAQPSSASSASTAAVAAN 325
 |||||
 RESULT 4
 ADD93650
 ID ADD93650 standard; protein; 431 AA.
 XX
 AC ADD93650;
 XX
 DT 29-JAN-2004 (first entry)
 XX
 DE Streptococcus mutans glucan binding protein-B.
 XX
 KW Glucan binding protein-B; GbpB; vaccine; anticaries; epitope; immunogen.
 XX
 OS Streptococcus mutans.
 XX
 PN WO2003075845-A2.
 XX
 PD 18-SEP-2003.
 XX
 PF 07-MAR-2003; 2003WO-US006962.
 XX
 PR 07-MAR-2002; 2002US-0363209P.
 PR 08-AUG-2002; 2002US-0402483P.
 XX
 PA (FORS-) FORSYTH INST.
 XX
 PI Smith DJ, Taubman MA;
 XX
 DR WPI; 2003-845091/78.
 DR GENBANK; AY046411.
 XX
 CC Composition useful as vaccines for dental caries comprises a fragment of
 PT a glucan binding protein-B binding to a major histocompatibility complex
 PT class II protein.
 XX
 PS Claim 5; Page 8; 49pp; English.
 XX
 CC The present sequence is the protein sequence of the glucan binding
 CC protein-B (GbpB) of Streptococcus mutans strain 3VF4. The sequence
 CC includes a number of human leucocyte antigen (HLA)-binding epitopes. The
 CC invention provides immunogenic compositions and vaccines for dental

CC caries. The compositions comprise major histocompatibility complex (MHC)
 CC class II protein-binding GbpB peptides, especially HLA-binding peptides,
 CC covalently linked with peptide subunits of a glucosyltransferase. The
 CC compositions are used in a claimed method of eliciting production of an
 CC antibody in a mammal. Diepitopic or multiepitopic polypeptides can be
 CC prepared synthetically or by recombinant DNA technology. Antibodies
 CC raised against MHC class II binding fragments of GbpB can be used in
 CC passive immunisation.
 XX
 SQ Sequence 431 AA;
 Query Match 100.0%; Score 90; DB 7; Length 431;
 Best Local Similarity 100.0%; Pred. No. 5.7e-05;
 Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 TATEAQPSSASSASTAAVAAN 20
 Db 306 TATEAQPSSASSASTAAVAAN 325
 |||||
 RESULT 5
 ADX37272
 ID ADX37272 standard; protein; 431 AA.
 XX
 AC ADX37272;
 XX
 DT 21-APR-2005 (first entry)
 XX
 DE Streptococcus mutant glucan binding protein B variant #1.
 XX
 KW immunogenicity; immune stimulation; glucan binding protein-B;
 KW microparticle; major histocompatibility complex; tooth disease.
 XX
 OS Streptococcus mutans.
 XX
 PN US2005031633-A1.
 XX
 PD 10-FEB-2005.
 XX
 PF 09-MAR-2004; 2004US-00797821.
 XX
 PR 13-APR-1998; 98US-0081550P.
 PR 08-JAN-1999; 99US-0115142P.
 PR 12-APR-1999; 99US-00290049.
 PR 07-MAR-2002; 2002US-0363209P.
 PR 08-AUG-2002; 2002US-0402483P.
 PR 07-MAR-2003; 2003US-00383930.
 XX
 PA (SMIT/) SMITH D J.
 PA (TAUB/) TAUBMAN M A.
 XX
 PI Smith DJ, Taubman MA;
 XX
 DR WPI; 2005-151644/16.
 XX
 PT New composition comprising a fragment of a glucan binding protein-B
 PT (GbpB) that binds to MHC class II protein, and a biocompatible
 PT microparticle, useful for producing an antibody (claimed) for immunizing
 PT mammals against dental caries.
 XX
 PS Claim 3; SEQ ID NO 29; 73pp; English.
 XX
 CC The invention relates to a composition comprising a fragment of a glucan
 CC binding protein-B (GbpB) and a biocompatible microparticle, where the
 CC fragment binds to a major histocompatibility complex (MHC) class II
 CC protein. The composition is useful for producing an antibody for
 CC immunizing mammals against dental caries. This sequence corresponds to a
 CC Streptococcus mutans GbpB protein of the invention.
 XX
 SQ Sequence 431 AA;
 Query Match 100.0%; Score 90; DB 9; Length 431;
 Best Local Similarity 100.0%; Pred. No. 5.7e-05;

Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TATEAQPSSASTAAVAAN 20
 Db 306 TATEAQPSSASTAAVAAN 325

RESULT 6
 ADX37273
 ID ADX37273 standard; protein; 431 AA.
 XX
 AC ADX37273;
 XX
 XX 21-APR-2005 (first entry)
 XX
 DE Streptococcus mutant glucan binding protein B variant #2.
 XX
 XX immunogenicity; immune stimulation; glucan binding protein-B;
 KW microparticle; major histocompatibility complex; tooth disease.
 KW
 XX Streptococcus mutans.
 OS
 XX Streptococcus mutans.
 PN US2005031633-A1.
 XX
 XX 10-FEB-2005.
 PD
 XX 09-MAR-2004; 2004US-00797821.
 PF
 XX 13-APR-1998; 98US-0081550P.
 PR
 PR 08-JAN-1999; 99US-0115142P.
 PR 12-APR-1999; 99US-00290049.
 PR 07-MAR-2002; 2002US-0363209P.
 PR 08-AUG-2002; 2002US-0402483P.
 PR 07-MAR-2003; 2003US-00363930.
 XX
 XX (SMIT/) SMITH D J.
 PA (TAUB/) TAUBMAN M A.
 XX
 XX Smith DJ, Taubman MA;
 PI
 XX WPI; 2005-151644/16.
 DR
 XX New composition comprising a fragment of a glucan binding protein-B
 PT (GbpB) that binds to MHC class II protein, and a biocompatible
 PT microparticle, useful for producing an antibody (claimed) for immunizing
 PT mammals against dental caries.
 XX
 XX Claim 3; SEQ ID NO 30; 73pp; English.
 PS
 XX The invention relates to a composition comprising a fragment of a glucan
 XX binding protein-B (GbpB) and a biocompatible microparticle, where the
 CC fragment binds to a major histocompatibility complex (MHC) class II
 CC protein. The composition is useful for producing an antibody for
 CC immunizing mammals against dental caries. This sequence corresponds to a
 CC Streptococcus mutans GbpB protein of the invention.
 XX
 XX Sequence 431 AA;
 SQ

Query Match 100.0%; Score 90; DB 9; Length 431;
 Best Local Similarity 100.0%; Pred. No. 5.7e-05;
 Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TATEAQPSSASTAAVAAN 20
 Db 306 TATEAQPSSASTAAVAAN 325

RESULT 7
 AEB91500
 ID AEB91500 standard; protein; 431 AA.
 XX
 AC AEB91500;
 XX

20-OCT-2005 (first entry)
 DT
 XX Microbial pathogen adhesin protein sequence, SEQ ID NO:210.
 DE
 XX algorithm; adhesin; pharmaceutical; vaccine; drug screening;
 KW bordetella pertussis infection; antibacterial; pneumonia;
 KW antiinflammatory; respiratory-gen.; gastric ulcer; antilulcer;
 KW gastrointestinal-gen.; urinary tract infection; antimicrobial; uropathic.
 XX
 XX Streptococcus mutans.
 OS
 XX WO2005076010-A2.
 PN
 XX 18-AUG-2005.
 PD
 XX 07-FEB-2005; 2005WO-IN000037.
 PF
 XX 06-FEB-2004; 2004IN-DE000173.
 PR
 PR 20-JUL-2004; 2004US-0589227P.
 XX
 XX (COUL) COUNCIL SCI & IND RES SOUTH AFRICA.
 XX
 XX Sachdeva G, Kumar K, Jain P, Brahmachari SK, Ramachandran S;
 PI
 XX WPI; 2005-597835/61.
 DR
 XX Computational method for identifying adhesin and adhesin like molecules,
 PT comprises computing sequence-based attributes of protein sequences using
 PT neural network software and training an artificial neural network.
 PT
 PT Claim 16; SEQ ID NO 210; 402pp; English.
 PS
 XX The present invention relates to a computational method (M1) for
 CC identifying adhesin and adhesin-like proteins, by computing the sequence-
 CC based attributes of protein sequences using five attribute modules of a
 CC neural network software, training an artificial neural network (ANN) for
 CC each of the computed five attributes, and identifying the adhesin and
 CC adhesin-like proteins having probability of being an adhesin (Pad) as
 CC equal or greater than 0.51. Also claimed is a set of 274 annotated genes
 CC encoding adhesin and adhesin-like proteins, having 274 fully defined 162-
 CC 1151 base pair (SEQ ID NO: 385-658) sequences; a set of 105 hypothetical
 CC genes encoding adhesin and adhesin-like proteins, having 105 fully
 CC defined 306-15876 base pairs (SEQ ID NO: 659-763); a set of 279 annotated
 CC adhesin and adhesin-like proteins, having 279 fully defined 53-3716 base
 CC pair (SEQ ID NO: 1-279); a set of 105 hypothetical adhesin and adhesin-
 CC like proteins, having 105 fully defined 106-5291 base pair (SEQ ID NO:
 CC 280-384) sequences; and a fully connected multilayer feed forward ANN (1)
 CC based on (M1). (M1) is useful for identifying adhesin and adhesin-like
 CC proteins, of therapeutic potential, and identifying and short-listing
 CC proteins for further testing in development of new vaccine formulations
 CC to eliminate diseases caused by various pathogenic organisms. (M1) is
 CC useful for identifying putative adhesins that are important in drug
 CC discovery and preventing therapeutics for whooping cough, pneumonia,
 CC gastric ulcer and urinary tract infections. (M1) identifies adhesins from
 CC distantly related organisms, and from bacteria belonging to a wide
 CC phylogenetic spectrum. (M1) is capable of predicting adhesive nature of
 CC unique proteins. The present sequence is a microbial pathogen adhesin
 CC protein sequence.
 XX
 XX Sequence 431 AA;
 SQ

Query Match 100.0%; Score 90; DB 9; Length 431;
 Best Local Similarity 100.0%; Pred. No. 5.7e-05;
 Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TATEAQPSSASTAAVAAN 20
 Db 306 TATEAQPSSASTAAVAAN 325

RESULT 8
 ADD93651
 ID ADD93651 standard; protein; 432 AA.

```

XX AC ADD93651;
XX DT 29-JAN-2004 (first entry)
XX DE Streptococcus mutans glucan binding protein-B.
XX KW Glucan binding protein-B; GbpB; vaccine; anticaries; epitope; immunogen.
XX OS Streptococcus mutans.
XX PN WO2003075845-A2.
XX PD 18-SEP-2003.
XX PF 07-MAR-2003; 2003WO-US0006962.
XX PR 07-MAR-2002; 2002US-0363209P.
XX PS 08-AUG-2002; 2002US-0402483P.
XX PA (FORS-) FORSYTH INST.
XX PI Smith DJ, Taubman MA;
XX PF WPI; 2003-845091/78.
XX PR GENBANK; AY046412.
XX PT Composition useful as vaccines for dental caries comprises a fragment of
XX PT a glucan binding protein-B binding to a major histocompatibility complex
XX PT class II protein.
XX PS Claim 5; Page 8; 49pp; English.
XX CC The present sequence is the protein sequence of the glucan binding
XX CC protein-B (GbpB) of Streptococcus mutans strain 15Jp2. The sequence
XX CC includes a number of human leucocyte antigen (HLA)-binding epitopes. The
XX CC invention provides immunogenic compositions and vaccines for dental
XX CC caries. The compositions comprise major histocompatibility complex (MHC)
XX CC class II protein-binding GbpB peptides, especially HLA-binding peptides,
XX CC covalently linked with peptide subunits of a glucosyltransferase. The
XX CC compositions are used in a claimed method of eliciting production of an
XX CC antibody in a mammal. Diepitopic or multiepitopic polypeptides can be
XX CC prepared synthetically or by recombinant DNA technology. Antibodies
XX CC raised against MHC class II binding fragments of GbpB can be used in
XX CC passive immunisation.
XX PS Sequence 432 AA;
XX CC
XX CC Query Match 93.3%; Score 84; DB 7; Length 432;
XX CC Best Local Similarity 95.0%; Pred. No. 0.00045;
XX CC Matches 19; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 TATAQPSASSASTAAVAAN 20
Db 307 TATAQPSASSASTAAVAAN 326
|||||
|||||

RESULT 9
ADD93652
ID ADD93652 standard; protein; 432 AA.
XX AC ADD93652;
XX DT 29-JAN-2004 (first entry)
XX DE Streptococcus mutans glucan binding protein-B.
XX KW Glucan binding protein-B; GbpB; vaccine; anticaries; epitope; immunogen.
XX OS Streptococcus mutans.
XX PN WO2003075845-A2.
XX PA (SMIT/) SMITH D J.

```

```

PD 18-SEP-2003.
XX PF 07-MAR-2003; 2003WO-US0006962.
XX PR 07-MAR-2002; 2002US-0363209P.
XX PS 08-AUG-2002; 2002US-0402483P.
XX PA (FORS-) FORSYTH INST.
XX PI Smith DJ, Taubman MA;
XX PF WPI; 2003-845091/78.
XX PR GENBANK; AY046413.
XX PT Composition useful as vaccines for dental caries comprises a fragment of
XX PT a glucan binding protein-B binding to a major histocompatibility complex
XX PT class II protein.
XX PS Claim 5; Page 8; 49pp; English.
XX CC The present sequence is the protein sequence of the glucan binding
XX CC protein-B (GbpB) of Streptococcus mutans strain 3SN1. The sequence
XX CC includes a number of human leucocyte antigen (HLA)-binding epitopes. The
XX CC invention provides immunogenic compositions and vaccines for dental
XX CC caries. The compositions comprise major histocompatibility complex (MHC)
XX CC class II protein-binding GbpB peptides, especially HLA-binding peptides,
XX CC covalently linked with peptide subunits of a glucosyltransferase. The
XX CC compositions are used in a claimed method of eliciting production of an
XX CC antibody in a mammal. Diepitopic or multiepitopic polypeptides can be
XX CC prepared synthetically or by recombinant DNA technology. Antibodies
XX CC raised against MHC class II binding fragments of GbpB can be used in
XX CC passive immunisation.
XX PS Sequence 432 AA;
XX CC
XX CC Query Match 93.3%; Score 84; DB 7; Length 432;
XX CC Best Local Similarity 95.0%; Pred. No. 0.00045;
XX CC Matches 19; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 TATAQPSASSASTAAVAAN 20
Db 307 TATAQPSASSASTAAVAAN 326
|||||
|||||

RESULT 10
ADX37274
ID ADX37274 standard; protein; 432 AA.
XX AC ADX37274;
XX DT 21-APR-2005 (first entry)
XX DE Streptococcus mutant glucan binding protein B variant #3.
XX KW immunogenicity; immune stimulation; glucan binding protein-B;
XX KW microparticle; major histocompatibility complex; tooth disease.
XX OS Streptococcus mutans.
XX PN US2005031633-A1.
XX PD 10-FEB-2005.
XX PF 09-MAR-2004; 2004US-00797821.
XX PR 13-APR-1998; 98US-0081550P.
XX PR 08-JAN-1999; 99US-0115142P.
XX PR 12-APR-1999; 99US-00290049.
XX PR 07-MAR-2002; 2002US-0363209P.
XX PR 08-AUG-2002; 2002US-0402483P.
XX PR 07-MAR-2003; 2003US-00383930.
XX PA (SMIT/) SMITH D J.

```



```

QY 1 TATEAOPSASSASTAAVAAN 20
Db 306 TATEAOPSASSASTAAVVTAN 325

RESULT 13
ADX37276
ID ADX37276 standard; protein; 431 AA.
XX AC ADX37276;
XX DT 21-APR-2005 (first entry)
XX DE Streptococcus mutant glucan binding protein B variant #5.
XX KW immunogenicity; immune stimulation; glucan binding protein-B;
XX KW microparticle; major histocompatibility complex; tooth disease.
XX OS Streptococcus mutans.
XX PN US2005031633-A1.
XX PD 10-FEB-2005.
XX PF 09-MAR-2004; 2004US-00797821.
XX PR 13-APR-1998; 98US-0081550P.
XX PR 08-JAN-1999; 99US-0115142P.
XX PR 12-APR-1999; 99US-00290049.
XX PR 07-MAR-2002; 2002US-0363209P.
XX PR 08-AUG-2002; 2002US-0402483P.
XX PR 07-MAR-2003; 2003US-00383930.
XX PA (SMIT/) SMITH D J.
XX PA (TAUB/) TAUBMAN M A.
XX PI Smith DJ, Taubman MA;
XX WPI; 2005-151644/16.
XX DE New composition comprising a fragment of a glucan binding protein-B
XX PT (GbpB) that binds to MHC class II protein, and a biocompatible
XX PT microparticle, useful for producing an antibody (claimed) for immunizing
XX PT mammals against dental caries.
XX PS Claim 3; SEQ ID NO 33; 73pp; English.
XX CC The invention relates to a composition comprising a fragment of a glucan
XX CC binding protein-B (GbpB) and a biocompatible microparticle, where the
XX CC fragment binds to a major histocompatibility complex (MHC) class II
XX CC protein. The composition is useful for producing an antibody for
XX CC immunizing mammals against dental caries. This sequence corresponds to a
XX CC Streptococcus mutans GbpB protein of the invention.
XX SQ Sequence 431 AA;

Query Match 91.1%; Score 82; DB 9; Length 431;
Best Local Similarity 90.0%; Pred. No. 0.00089;
Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 TATEAOPSASSASTAAVAAN 20
Db 306 TATEAOPSASSASTAAVVTAN 325

RESULT 14
ADD93638
ID ADD93638 standard; peptide; 20 AA.
XX AC ADD93638;
XX DT 29-JAN-2004 (first entry)
XX PD 10-FEB-2005.
XX

Streptococcus mutans glucan binding protein-B peptide fragment.
Glucan binding protein-B; GbpB; vaccine; anticaries; epitope; immunogen.
Streptococcus mutans.
WO2003075845-A2.
18-SEP-2003.
07-MAR-2003; 2003WO-US006962.
07-MAR-2002; 2002US-0363209P.
08-AUG-2002; 2002US-0402483P.
(FORS-) FORSYTH INST.
Smith DJ, Taubman MA;
WPI; 2003-845091/78.
Composition useful as vaccines for dental caries comprises a fragment of
a glucan binding protein-B binding to a major histocompatibility complex
class II protein.
Claim 11; Page 10; 49pp; English.
The present sequence is that of a peptide comprising amino acid residues
311-330 of the glucan binding protein-B (GbpB) of Streptococcus mutans
strain SK32 ADD93649. The peptide binds to a major histocompatibility
complex (MHC) class II protein. It was identified as a potential B cell
epitope using a matrix-based algorithm for epitope prediction, which was
used to search the primary amino acid sequence of gbpB for known MHC
class II binding motifs. The peptide can be used in immunogenic
compositions and subunit vaccines for dental caries. These comprise MHC
class II protein-binding GbpB peptides covalently linked with peptide
subunits (preferably from the catalytic domain) of a glucosyltransferase.
The compositions are used in a claimed method of eliciting production of
an antibody in a mammal. Diepitopic or multipitopic polypeptides can be
prepared synthetically or by recombinant DNA technology. Antibodies
raised against MHC class II binding fragments of GbpB can be used in
passive immunisation.
SQ Sequence 20 AA;

Query Match 74.4%; Score 67; DB 7; Length 20;
Best Local Similarity 100.0%; Pred. No. 0.0054;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 6 QPSASSASTAAVAAN 20
Db 1 QPSASSASTAAVAAN 15

RESULT 15
ADX37261
ID ADX37261 standard; peptide; 20 AA.
XX AC ADX37261;
XX DT 21-APR-2005 (first entry)
XX DE Streptococcus mutant glucan binding protein B peptide #18.
XX KW immunogenicity; immune stimulation; glucan binding protein-B;
XX KW microparticle; major histocompatibility complex; tooth disease.
XX OS Streptococcus mutans.
XX PN US2005031633-A1.
XX PD 10-FEB-2005.
XX

```

PF 09-MAR-2004; 2004US-00797821.
XX
PR 13-APR-1998; 98US-0081550P.
PR 08-JAN-1999; 99US-0115142P.
PR 12-APR-1999; 99US-00290049.
PR 07-MAR-2002; 2002US-0363209P.
PR 08-AUG-2002; 2002US-0402483P.
PR 07-MAR-2003; 2003US-00383930.
XX
PA (SMIT/) SMITH D J.
PA (TAUB/) TAUBMAN M A.
XX
PI Smith DJ, Taubman MA;
XX
DR WPI; 2005-151644/16.
XX
PT New composition comprising a fragment of a glucan binding protein-B
PT (GbpB) that binds to MHC class II protein, and a biocompatible
PT microparticle, useful for producing an antibody (claimed) for immunizing
PT mammals against dental caries.
XX
PS Claim 4; SEQ ID NO 18; 73pp; English.
XX
CC The invention relates to a composition comprising a fragment of a glucan
CC binding protein-B (GbpB) and a biocompatible microparticle, where the
CC fragment binds to a major histocompatibility complex (MHC) class II
CC protein. The composition is useful for producing an antibody for
CC immunizing mammals against dental caries. This sequence corresponds to a
CC Streptococcus mutans GbpB-derived peptide of the invention.
XX
SQ Sequence 20 AA;

Query Match 74.4%; Score 67; DB 9; Length 20;
Best Local Similarity 100.0%; Pred. No. 0.0054;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 6 QPSASASTAAVAAN 20
| | | | | | | | | |
Db 1 QPSASASTAAVAAN 15

Search completed: January 25, 2006, 19:01:41
Job time : 69.2 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2006 Compugen Ltd.

OM protein - protein search, using sw model

Run on: January 25, 2006, 18:48:55 ; Search time 9.1 Seconds
(without alignments)
211.465 Million cell updates/sec

Title: US-10-797-821-3
Perfect score: 90
Sequence: 1 TATEAQPSASSASTAAVAAN 20

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283416 seqs, 96216763 residues

Total number of hits satisfying chosen parameters: 283416

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : PIR 80:*
1: pir1:*
2: pir2:*
3: pir3:*
4: pir4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	48	53.3	172	2 H86837	hypothetical prote
2	48	53.3	828	2 C88402	protein H05C05.1 [
3	47	52.2	384	2 D86448	F5D14.12 protein -
4	47	52.2	402	2 A86194	hypothetical prote
5	47	52.2	494	2 A42170	zinc finger protei
6	47	52.2	497	2 JC5076	myc-associated zin
7	47	52.2	644	2 S39356	transcription fact
8	47	52.2	825	1 EDBEXD	immediate-early pr
9	46	51.1	125	2 G72716	hypothetical prote
10	46	51.1	347	2 JC7178	chitinase (EC.3.2.
11	46	51.1	364	2 T02801	probable membrane
12	46	51.1	372	2 T45524	regulatory protein
13	46	51.1	1224	2 T14007	microtubule-associ
14	45	50.0	1446	1 A45344	immediate-early pr
15	44	48.9	317	2 T00500	probable elicitor
16	44	48.9	331	2 B47236	zinc-finger protei
17	44	48.9	439	1 VHBPCL	minor capsid prote
18	44	48.9	439	2 A99833	probable capsid pr
19	44	48.9	439	2 C85690	probable capsid pr
20	44	48.9	439	2 H90900	probable minor cap
21	44	48.9	477	2 A47236	zinc-finger protei
22	44	48.9	503	2 S63257	probable membrane
23	44	48.9	724	2 S57604	probable membrane
24	44	48.9	797	1 VGBEX1	glycoprotein X pre
25	44	48.9	866	2 T45462	membrane glycoprot
26	44	48.9	867	2 T45463	membrane glycoprot
27	44	48.9	1156	2 T34852	probable secreted
28	44	48.9	1348	2 S27812	probable epidermal
29	44	48.9	1348	2 A43917	probable epidermal

30	44	48.9	1355	2 S40022	spalt protein - fr
31	44	48.9	1799	1 S44920	ZK688.5 protein -
32	44	48.9	2061	2 T13751	transcription fact
33	43	47.8	85	1 FDFL4W	antifreeze protein
34	43	47.8	91	2 A22592	antifreeze protein
35	43	47.8	141	2 H83011	translocation prot
36	43	47.8	285	1 A37757	acetolactate decar
37	43	47.8	363	2 F87754	protein C43E11.1 [
38	43	47.8	453	2 H87373	hypothetical prote
39	43	47.8	507	2 T44768	antifreeze glycope
40	43	47.8	547	2 H83018	dihydroliipoamide a
41	43	47.8	601	1 S11777	hydrogenase (EC 1.
42	43	47.8	656	2 T01573	earl Protein - mai
43	43	47.8	785	2 T03491	hemagglutinin, pha
44	43	47.8	822	2 A38420	antifreeze glycopr
45	43	47.8	1013	2 T46422	hypothetical prote

ALIGNMENTS

RESULT 1

H86837
hypothetical protein yrgA [imported] - Lactococcus lactis subsp. lactis (strain IL1403)
C:Species: Lactococcus lactis subsp. lactis
C:Date: 23-Mar-2001 #sequence_revision 23-Mar-2001 #text_change 09-Jul-2004
C:Accession: H86837
R:Bolotin, A.; Wincker, P.; Mauder, S.; Jaillon, O.; Malarre, K.; Weissbach, J.; Ehrlich
Genome Res. 11, 731-753, 2001
A:Title: The complete genome sequence of the lactic acid bacterium Lactococcus lactis s
A:Reference number: A86625; MUID:21235186; PMID:11337471
A:Accession: H86837
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-172 <STO>
A:Cross-references: UNIPROT:Q9CEY0; UNIPARC:UPI000000D4A2; GB:AE005176; PID:gl2724721;
A:Experimental source: strain IL1403
C:Genetics:
A:Gene: yrgA

Query Match 53.3%; Score 48; DB 2; Length 172;
Best Local Similarity 50.0%; Pred. No. 7.6;
Matches 10; Conservative 6; Mismatches 4; Indels 0; Gaps 0;

QY 1 TATEAQPSASSASTAAVAAN 20
::||| ||||| |:::
Db 95 SSTEATPPASSSTVASSSS 114

RESULT 2

C88402
protein H05C05.1 [imported] - Caenorhabditis elegans
C:Species: Caenorhabditis elegans
C:Date: 10-May-2001 #sequence_revision 10-May-2001 #text_change 09-Jul-2004
C:Accession: C88402
R:Anonymous, The C. elegans Sequencing Consortium.
Science 282, 2012-2018, 1998
A:Title: Genome sequence of the nematode C. elegans: a platform for investigating biolo
A:Reference number: A75000; MUID:99069613; PMID:9851916
A:Note: see websites genome.wustl.edu/gsc/C_elegans/ and www.sanger.ac.uk/Projects/C_el
A:Note: published errata appeared in Science 283, 35, 1999; Science 283, 2103, 1999; an
A:Accession: C88402
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-828 <STO>
A:Cross-references: UNIPROT:Q9TXU2; UNIPARC:UPI000007CE2B; GB:chr_III; PIDN:AAC68790.1;
C:Genetics:
A:Gene: H05C05.1
A:Map position: 3

Query Match 53.3%; Score 48; DB 2; Length 828;
Best Local Similarity 57.9%; Pred. No. 35;
Matches 11; Conservative 2; Mismatches 6; Indels 0; Gaps 0;

QY 2 ATEAQPSSASSASTAAVAAN 20
| | | | |
Db 766 AASAQFSESSSATAASGDN 784

RESULT 3
D86448
F5D14.12 protein - Arabidopsis thaliana
C:Species: Arabidopsis thaliana (mouse-ear cress)
C:Date: 02-Mar-2001 #sequence_revision 02-Mar-2001 #text_change 09-Jul-2004
C:Accession: D86448
R:Theologis, A.; Ecker, J.R.; Palm, C.J.; Federspiel, N.A.; Kaul, S.; White, O.; Alonso,
Chin, C.W.; Chung, M.K.; Conn, L.; Conway, A.B.; Creasy, T.H.; Dewar, K.;
ansen, N.F.; Hughes, B.; Huizar, L.
Nature 408, 816-820, 2000
A:Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kim, C.
C.A.; Li, J.H.; Li, Y.; Lin, X.; Liu, S.X.; Liu, Z.A.; Lueros, J.S.; Maitl, R.; Marziali,
Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.
A:Authors: Salzberg, S.L.; Schwartz, J.R.; Shinn, P.; Southwick, A.M.; Sun, H.; Tallon,
ker, M.; Wu, D.; Yu, G.; Fraser, C.M.; Venter, J.C.; Davis, R.W.
A>Title: Sequence and analysis of chromosome 1 of the plant Arabidopsis.
A:Reference number: A86141; PMID:11130712
A:Accession: D86448
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-384 <STO>
A:Cross-references: UNIPROT:Q9LQM3; UNIPARC:UPI00000AC169; GB:AE005172; NID:g8920610; PT:
C:Genetics:
A:Map position: 1

Query Match 52.2%; Score 47; DB 2; Length 384;
Best Local Similarity 50.0%; Pred. No. 23;
Matches 9; Conservative 4; Mismatches 5; Indels 0; Gaps 0;

QY 3 TEAQPSASSASTAAVAAN 20
| | | | | : | |
Db 65 TRGSPSSSATTTSAASN 82

RESULT 4
A86194
hypothetical protein [imported] - Arabidopsis thaliana
C:Species: Arabidopsis thaliana (mouse-ear cress)
C:Date: 02-Mar-2001 #sequence_revision 02-Mar-2001 #text_change 09-Jul-2004
C:Accession: A86194
R:Theologis, A.; Ecker, J.R.; Palm, C.J.; Federspiel, N.A.; Kaul, S.; White, O.; Alonso,
Chin, C.W.; Chung, M.K.; Conn, L.; Conway, A.B.; Creasy, T.H.; Dewar, K.;
ansen, N.F.; Hughes, B.; Huizar, L.
Nature 408, 816-820, 2000
A:Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kim, C.
C.A.; Li, J.H.; Li, Y.; Lin, X.; Liu, S.X.; Liu, Z.A.; Lueros, J.S.; Maitl, R.; Marziali,
Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.
A:Authors: Salzberg, S.L.; Schwartz, J.R.; Shinn, P.; Southwick, A.M.; Sun, H.; Tallon,
ker, M.; Wu, D.; Yu, G.; Fraser, C.M.; Venter, J.C.; Davis, R.W.
A>Title: Sequence and analysis of chromosome 1 of the plant Arabidopsis.
A:Reference number: A86141; PMID:11130712
A:Accession: A86194
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-402 <STO>
A:Cross-references: UNIPROT:Q9WA35; UNIPARC:UPI00000A87F1; GB:AE005172; NID:g6850320; PT:
C:Genetics:
A:Map position: 1

```

RESULT 5
A42170
N;zinc finger protein MAZ - human (fragment)
N;Alternate names: MYC-associated zinc finger protein MAZ; zinc finger protein ZF87
C;Species: Homo sapiens (man)
C;Date: 03-Mar-1994 #sequence_revision 03-Mar-1994 #text_change 09-Jul-2004
C;Accession: A42170; A46153
R;Pyrc, J.J.; Moberg, K.H.; Hall, D.J.
Biochemistry 31, 4102-4110, 1992
A;Title: Isolation of a novel cDNA encoding a zinc-finger protein that binds to two sites
A;Reference number: A42170; MUID:92232709; PMID:1567856
A;Accession: A42170
A;Status: not compared with conceptual translation
A;Molecule type: mRNA
A;Residues: 1-494 <PYR>
A;Cross-references: UNIPROT:P56270; UNIPARC:UPI000017C427; GB:J05371
A;Note: it is uncertain whether Met-18 is the initiator or whether translation is initiated
R;Bossone, S.A.; Aesselin, C.; Patel, A.J.; Marcu, K.B.
Proc. Natl. Acad. Sci. U.S.A. 89, 7452-7456, 1992
A;Title: MAZ, a zinc finger protein, binds to c-MYC and C2 gene sequences regulating transcription
A;Reference number: A46153; MUID:92366479; PMID:1502157
A;Accession: A46153
A;Molecule type: mRNA
A;Residues: 18-417, 'L', 419-494 <BOS>
A;Cross-references: UNIPARC:UPI000012BCF8; GB:M94046
A;Experimental source: HeLa cells
A;Note: sequence extracted from NCBI backbone (NCBIN:110666, NCBI:P:110667)
C;Keywords: DNA binding; zinc finger
F;113-125/Region: alanine-rich
F;174-183/Region: alanine-rich
F;207-230/Region: zinc finger
F;296-318/Region: zinc finger
F;324-346/Region: zinc finger
F;354-368/Region: zinc finger
F;373-405/Region: zinc finger
F;409-430/Region: zinc finger
F;452-468/Region: alanine-rich

Query Match 52.2%; Score 47; DB 2; Length 494;
Best Local Similarity 61.1%; Pred. No. 29;
Matches 11; Conservative 2; Mismatches 5; Indels 0; Gaps 0;

QY 2 ATEAOPSASSASTAAVAA 19
    ||| |||:: |||
DB 161 AAERAPPASANTIAAAA 178

RESULT 6
JC5076
myc-associated zinc-finger protein - human
N;Alternate names: MAZ protein
C;Species: Homo sapiens (man)
C;Date: 31-Jan-1997 #sequence_revision 31-Jan-1997 #text_change 05-Nov-1999
C;Accession: JC5076
R;Tsutsui, H.; Sakatsume, O.; Itakura, K.; Yokoyama, K.K.
Biochem. Biophys. Res. Commun. 226, 801-809, 1996
A;Title: Members of the MAZ family: A novel cDNA clone for MAZ from human pancreatic islet
A;Reference number: JC5076; MUID:96428591; PMID:8831693
A;Accession: JC5076
A;Molecule type: mRNA
A;Residues: 1-497 <TSU>
A;Cross-references: UNIPARC:UPI0000163B39; DDBJ:D85131; NID:gi752741; PIDN:BAA12728.1;
A;Experimental source: pancreatic islet
C;Comment: This protein plays a role in the control of transcriptional initiation of genes
and between the introns of the mouse gene for immunoglobulin M-D.
C;Keywords: phosphoprotein; zinc finger
F;146,204,480/Binding site: phosphate (Ser) (covalent) (by casein kinase II) #status predicted
F;349/Binding site: phosphate (Tyr) (covalent) #status predicted

Query Match 52.2%; Score 47; DB 2; Length 497;
Best Local Similarity 61.1%; Pred. No. 29;
Matches 11; Conservative 2; Mismatches 5; Indels 0; Gaps 0;

```


QY 2 ATEAOPSASSASTAAVAA 19
Db 170 AAEAPASPASATIAAAAA 187

RESULT 7
S39356
transcription factor btd - fruit fly (*Drosophila* sp.)
C:Species: *Drosophila* sp.
C:Date: 18-Feb-1994 #sequence_revision 01-Dec-1995 #text_change 07-May-1999
C:Accession: S39356
R:Wimmer, B.A.; Jaekle, H.; Pfeifle, C.; Cohen, S.M.
Nature 366, 690-694, 1993
A:Title: A *Drosophila* homologue of human Sp1 is a head-specific segmentation gene.
A:Reference number: S39356; MUID:94081952; PMID:8259212
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-644 <WIM>
A:Cross-references: UNIPARC:UPI0000124C17; EMBL:Z29361; NID:g441284
A:Gene: FlyBase:btd
A:Cross-references: FlyBase:FBgn0000233
A:Introns: 245/2

Query Match 52.2%; Score 47; DB 2; Length 644;
Best Local Similarity 58.8%; Pred. No. 38;
Matches 10; Conservative 3; Mismatches 4; Indels 0; Gaps 0;

QY 3 TEAOPSASSASTAAVAA 19
Db 197 TSSPSSSAASAAAAA 213

RESULT 8
EBEXD
immediate-early protein RL2 - human herpesvirus 2 (strain HG52)
N:Alternate names: RL2 protein
C:Species: human herpesvirus 2
A:Note: host *Homo sapiens* (man)
C:Date: 31-Dec-1992 #sequence_revision 31-Dec-1992 #text_change 09-Jul-2004
C:Accession: JQ1501
R:McGeoch, D.J.; Cunningham, C.; McIntyre, G.; Dolan, A.
J. Gen. Virol. 72, 3057-3075, 1991
A:Title: Comparative sequence analysis of the long repeat regions and adjoining parts of
A:Reference number: JQ1494; MUID:92113549; PMID:1662697
A:Accession: JQ1501
A:Molecule type: DNA
A:Residues: 1-825 <MCG>
A:Cross-references: UNIPROT:P28284; UNIPARC:UPI000012D179; GB:D10471; DBJ:D01128; NID:9
C:Genetics:
A:Gene: RL2
A:Introns: 25/3; 252/1
C:Superfamily: herpesvirus immediate-early protein IB110; RING finger homology
C:Keywords: DNA binding; immediate-early protein; tandem repeat; transcription regulation
F:122-172/Domain: RING finger homology <RING>
F:126-166/Region: zinc finger C3HC4 motif
F:589-623/Region: 5-residue repeats (A-S-S-S-S)

Query Match 52.2%; Score 47; DB 1; Length 825;
Best Local Similarity 63.2%; Pred. No. 48;
Matches 12; Conservative 2; Mismatches 5; Indels 0; Gaps 0;

QY 1 TATEAOPSASSASTAAVAA 19
Db 570 SAGAAPPSPSSQAATAA 588

RESULT 9
G72716
hypothetical protein APE1150 - *Aeropyrum pernix* (strain K1)
C:Species: *Aeropyrum pernix*

C:Date: 20-Aug-1999 #sequence_revision 20-Aug-1999 #text_change 09-Jul-2004
C:Accession: G72716
R:Kawarabayashi, Y.; Hino, Y.; Horikawa, H.; Yamazaki, S.; Haikawa, Y.; Jin-no, K.; Taka
awa, H.; Takamiya, M.; Masuda, S.; Funahashi, T.; Tanaka, T.; Kudoh, Y.; Yamazaki, J.;
DNA Res. 6, 83-101, 1999
A:Title: Complete genome sequence of an aerobic hyper-thermophilic Crenarchaeon, *Aeropy*
A:Reference number: A72450; MUID:99310339; PMID:10382966
A:Accession: G72716
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-125 <KAW>
A:Cross-references: UNIPROT:Q9YCW2; UNIPARC:UPI000005DE24; DBJ:AP0000060; NID:g5104188;
A:Experimental source: strain K1
C:Genetics:
A:Gene: APE1150

Query Match 51.1%; Score 46; DB 2; Length 125;
Best Local Similarity 58.8%; Pred. No. 11;
Matches 10; Conservative 4; Mismatches 3; Indels 0; Gaps 0;

QY 2 ATEAOPSASSASTAAVAA 18
Db 85 ATAAPPSTASTLSLS 101

RESULT 10
JC7178
chitinase (EC 3.2.1.14) - Streptomyces thermoviolaceus
C:Species: Streptomyces thermoviolaceus
C:Date: 04-Mar-2000 #sequence_revision 04-Mar-2000 #text_change 09-Jul-2004
C:Accession: JC7178
R:Tsujibo, H.; Hatanio, N.; Endo, H.; Miyamoto, K.; Inamori, Y.
Biosci. Biotechnol. Biochem. 64, 96-102, 2000
A:Title: Purification and characterization of a thermostable chitinase from Streptomyces
A:Reference number: JC7178; MUID:20169631; PMID:10705453
A:Accession: JC7178
A:Molecule type: DNA
A:Residues: 1-347 <TSU>
A:Cross-references: UNIPROT:Q9RHU3; UNIPARC:UPI000017AE40; DBJ:AB016844
A:Experimental source: strain OPC-520
C:Comment: This enzyme, involved in chitin degradation system, hydrolyzing chitin by an
he conversion of insoluble chitin to its monomeric component.
C:Genetics:
A:Gene: chi30
A:Start codon: GTG
C:Keywords: chitin biosynthesis; glycosidase; hydrolase
F:50-347/Product: 30K chitinase #status predicted <MAT>

Query Match 51.1%; Score 46; DB 2; Length 347;
Best Local Similarity 62.5%; Pred. No. 29;
Matches 10; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

QY 1 TATEAOPSASSASTAA 16
Db 37 TPTSAQPAAVTASSAA 52

RESULT 11
T02801
probable membrane protein L549.13 [imported] - Leishmania major (strain Friedlin)
C:Species: Leishmania major
C:Date: 24-Mar-1999 #sequence_revision 24-Mar-1999 #text_change 05-Oct-2004
C:Accession: E81456; T02801
R:Myler, P.J.; Audleman, L.; DeVos, T.; Hixson, G.; Kiser, P.; Lemley, C.; Magness, C.;
Proc. Natl. Acad. Sci. U.S.A. 96, 2902-2906, 1999
A:Title: Leishmania major Friedlin chromosome 1 has an unusual distribution of protein-
A:Reference number: A81455; MUID:99178987; PMID:10077609
A:Accession: E81456
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-364 <PYL>
A:Cross-references: UNIPROT:O60974; UNIPARC:UPI000007B776; GB:AE001274; NID:g3264850; P
A:Experimental source: strain MHOM/IL/81/Friedlin

C;Genetics:

A;Gene: L549.13

A;Map position: 1

C;Superfamily: Leishmania major probable membrane protein L549.13

C;Keywords: transmembrane protein

Query Match 51.1%; Score 46; DB 2; Length 364;

Best Local Similarity 83.3%; Pred. No. 30;

Matches 10; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 6 OPSASSASTAAV 17

:|||||:|||||

Db 220 KPSASSATAAV 231

RESULT 12

T45524

regulatory protein rim101 homolog [imported] - yeast (Kluyveromyces marxianus var. lactis

C;Species: Kluyveromyces marxianus var. lactis, Candida spharctica

C;Date: 31-Jan-2000 #sequence_revision 31-Jan-2000 #text_change 09-Jul-2004

C;Accession: T45524

R;Bao, W.G.; Fukuhara, H.

submitted to the EMBL Data Library, July 1999

A;Description: The ubiquitin-encoding genes of Kluyveromyces lactis.

A;Reference number: Z23000

A;Accession: T45524

A;Status: preliminary; translated from GB/EMBL/DBJ

A;Molecule type: DNA

A;Residues: 1-372 <BAO>

A;Cross-references: UNIPROT:Q9Y850; UNIPARC:UPI000006C4A8; EMBL:AJ243800; PIDN:CAB50896.

A;Experimental source: strain 2359/152

C;Genetics:

A;Gene: rim101

Query Match 51.1%; Score 46; DB 2; Length 372;

Best Local Similarity 71.4%; Pred. No. 31;

Matches 10; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 7 PSASSASTAAVAN 20

|||||:|||||

Db 279 PSASSASTAVSPN 292

RESULT 13

T14007

microtubule-associated protein 4 - African clawed frog

C;Species: Xenopus laevis (African clawed frog)

C;Date: 20-Sep-1999 #sequence_revision 20-Sep-1999 #text_change 09-Jul-2004

C;Accession: T14007

R;Shiina, N.; Teukita, S.

Mol. Biol. Cell 10, 597-608, 1999

A;Title: Mutations at phosphorylation sites of Xenopus microtubule-associated protein 4

A;Reference number: Z17855; MUID:99169009; PMID:10069806

A;Accession: T14007

A;Status: preliminary; translated from GB/EMBL/DBJ

A;Molecule type: mRNA

A;Residues: 1-1224 <SHI>

A;Cross-references: UNIPROT:Q9YI90; UNIPARC:UPI000000FDDE8; EMBL:AB021705; NID:dl1252552;

C;Keywords: microtubule binding

Query Match 51.1%; Score 46; DB 2; Length 1224;

Best Local Similarity 52.6%; Pred. No. 97;

Matches 10; Conservative 3; Mismatches 6; Indels 0; Gaps 0;

QY 1 TATEQPSASSASTAAVAA 19

:|||||:|||||

Db 846 TSVKASFPVASKSTATTAA 864

RESULT 14

A45344

immediate-early protein - suid herpesvirus 1 (strain Kaplan)

C;Species: suid herpesvirus 1

C;Date: 30-Sep-1993 #sequence_revision 30-Sep-1993 #text_change 09-Jul-2004

C;Accession: A45344

R;Vicek, C.; Kozmik, Z.; Paces, V.; Schirm, S.; Schwyzer, M.

Virology 179, 365-377, 1990

A;Title: Pseudorabies virus immediate-early gene overlaps with an oppositely oriented of

A;Reference number: A45344; MUID:91021039; PMID:2171211

A;Accession: A45344

A;Status: translation not shown

A;Molecule type: DNA

A;Residues: 1-1446 <VLC>

A;Cross-references: UNIPROT:P33479; UNIPARC:UPI000012D219; GB:M34651; NID:g334070; PIDN:

C;Superfamily: herpesvirus immediate-early protein IE175

C;Keywords: DNA binding; early protein; transcription regulation

Query Match 50.0%; Score 45; DB 1; Length 1446;

Best Local Similarity 55.6%; Pred. No. 1.6e+02;

Matches 10; Conservative 4; Mismatches 4; Indels 0; Gaps 0;

QY 2 ATEQPSASSASTAAVAA 19

:|||||:|||||

Db 424 AAMPSPASSASSASAAAS 441

RESULT 15

T00500

probable elicitor response element-binding protein WRKY3 - Arabidopsis thaliana

C;Species: Arabidopsis thaliana (mouse-ear cress)

C;Date: 01-Feb-1999 #sequence_revision 01-Feb-1999 #text_change 31-Dec-2004

C;Accession: T00500; C84623

R;Rounsley, S.D.; Lin, X.; Ketchum, K.A.; Crosby, M.L.; Brandon, R.C.; Sykes, S.M.; Kaul

submitted to the EMBL Data Library, November 1997

A;Description: Arabidopsis thaliana chromosome II BAC T20D16 genomic sequence.

A;Reference number: Z14164

A;Accession: T00500

A;Status: translated from GB/EMBL/DBJ

A;Molecule type: DNA

A;Residues: 1-317 <ROU>

A;Cross-references: UNIPROT:Q22176; UNIPARC:UPI0000138F7D; EMBL:AC002391; NID:g2642427;

A;Experimental source: cultivar Columbia

R;Lin, X.; Kaul, S.; Rounsley, S.D.; Shea, T.P.; Benito, M.I.; Town, C.D.; Fujii, C.Y.;

M.; Koo, H.; Moffat, K.S.; Cronin, L.A.; Shen, M.; VanAken, S.E.; Umayam, L.; Tallon, L.

euss, D.; Nierman, W.C.; White, O.; Eisen, J.A.; Salzberg, S.L.; Fraser, C.M.; Venter,

Nature 402, 761-768, 1999

A;Title: Sequence and analysis of chromosome 2 of the plant Arabidopsis thaliana.

A;Reference number: A84420; MUID:20083487; PMID:10617197

A;Accession: C84623

A;Status: preliminary

A;Molecule type: DNA

A;Residues: 1-317 <STO>

A;Cross-references: UNIPARC:UPI0000138F7D; GB:AE002093; NID:g2642432; PIDN:AAB87100.1;

C;Genetics:

A;Gene: WRKY3; ATSP:T20D16.5; At2g23320

A;Map position: 2

A;Introns: 218/2; 260/2

C;Superfamily: DNA-binding protein WRKY3

Query Match 48.9%; Score 44; DB 2; Length 317;

Best Local Similarity 71.4%; Pred. No. 51;

Matches 10; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 6 QPSASSASTAAVAA 19

|||||:|||||

Db 48 QPSSSSASASASAA 61

Search completed: January 25, 2006, 19:15:15

Job time : 9.1 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2006 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: January 25, 2006, 18:48:24 ; Search time 70.3 Seconds
(without alignments)
200.719 Million cell updates/sec

Title: US-10-797-821-3

Perfect score: 90

Sequence: 1 TATEAQPSASSASTAAVAAN 20

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 2166443 seqs, 705528306 residues

Total number of hits satisfying chosen parameters: 2166443

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : UniProt 05.80:*

1: uniprot_sprot:*

2: uniprot_trembl:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	90	100.0	431	Q938V3_STRMU	Q938v3 streptococc
2	90	100.0	431	Q9AG98_STRMU	Q9ag98 streptococc
3	90	100.0	431	Q8DWM3_STRMU	Q8dwm3 streptococc
4	84	93.3	432	Q938V1_STRMU	Q938v1 streptococc
5	84	93.3	432	Q938V2_STRMU	Q938v2 streptococc
6	82	91.1	431	Q938V0_STRMU	Q938v0 streptococc
7	55	61.1	166	Q84JT2_ORYSA	Q84jt2 oryza sativ
8	55	61.1	291	Q9VKU4_DROME	Q9vku4 drosophila
9	52	57.8	700	Q6YXX1_ORYSA	Q6yxx1 oryza sativ
10	52	57.8	1430	Q4NTX6_9DELT	Q4ntx6 anaeromyxob
11	51	56.7	1229	Q68H99_PIG	Q68h99 sus scrofa
12	51	56.7	1312	Q9UL13_LEIMA	Q9ul13 leishmania
13	51	56.7	1330	Q4WDG9_ASPFU	Q4wdg9 aspergillus
14	51	56.7	2193	Q4PFY0_USTMA	Q4pfy0 ustilago ma
15	50	55.6	481	Q5LR10_SILPO	Q5lri10 silicibacte
16	50	55.6	545	Q91TR2_TUHV1	Q91tr2 tupaiid her
17	50	55.6	1068	Q4QG70_LEIMA	Q4qg70 leishmania
18	50	55.6	2202	Q7SEAL_NEUCR	Q7seal neurospora
19	49	54.4	272	Q6H8X7_STRSU	Q6h8x7 streptococc
20	49	54.4	320	Q8RKT9_SERMA	Q8rkt9 serratia ma
21	49	54.4	601	Q63JW1_BURPS	Q63jw1 burkholderi
22	49	54.4	1090	Q6CMU7_KLULA	Q6cmu7 kluyvetomyc
23	49	54.4	1783	Q4Q5V5_LEIMA	Q4q5v5 leishmania
24	48	53.3	104	Q5S76_CANAL	Q5s76 candida alb
25	48	53.3	172	Q9CEY0_LACLA	Q9cey0 lactococcu
26	48	53.3	265	Q6H8W9_STRSU	Q6h8w9 streptococc
27	48	53.3	265	Q6H8X5_STRSU	Q6h8x5 streptococc
28	48	53.3	265	Q6H8X6_STRSU	Q6h8x6 streptococc
29	48	53.3	269	Q6H8X1_STRSU	Q6h8x1 streptococc
30	48	53.3	272	Q6H8X4_STRSU	Q6h8x4 streptococc
31	48	53.3	332	Q5BEZ6_EMENI	Q5bez6 aspergillus

32	48	53.3	408	2	Q5LQI7_SILPO	Q5lqi7 silicibacte
33	48	53.3	441	2	Q9NWF8_HUMAN	Q9nwf8 homo sapien
34	48	53.3	446	2	Q8IXM5_HUMAN	Q8ixm5 homo sapien
35	48	53.3	447	2	Q5MNZ9_HUMAN	Q5mnz9 homo sapien
36	48	53.3	658	2	Q91TV8_TUHV1	Q91tv8 tupaiid her
37	48	53.3	722	2	Q7XRF9_ORYSA	Q7xrf9 oryza sativ
38	48	53.3	820	2	Q9TXU2_CAREL	Q9txu2 caenorhabdi
39	48	53.3	1151	2	Q6H8Y4_STRSU	Q6h8y4 streptococc
40	48	53.3	1153	2	Q6H8Y1_STRSU	Q6h8y1 streptococc
41	48	53.3	1417	2	Q5CV93_CRYPV	Q5cv93 cryptospori
42	48	53.3	2205	2	Q4RTF3_TETNG	Q4rtf3 tetraodon n
43	48	53.3	2706	2	Q4RGM4_TETNG	Q4rgm4 tetraodon n
44	47.5	52.8	1357	2	Q4Q825_LEIMA	Q4q825 leishmania
45	47	52.2	121	2	Q4KU35_PLAMV	Q4ku35 nandina mos

ALIGNMENTS

RESULT 1
Q938V3_STRMU
ID Q938V3_STRMU PRELIMINARY; PRT; 431 AA.
AC Q938V3;
DT 01-DEC-2001 (Tremblrel. 19, Created)
DT 01-DEC-2001 (Tremblrel. 19, Last sequence update)
DT 01-MAR-2004 (Tremblrel. 26, Last annotation update)
DE Glucan-binding protein B.
OS Streptococcus mutans.
OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
OC Streptococcus.
OX NCBI_TaxID=1309;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=5J32;
RA Jin S., Duncan M.J., Taubman M.A., Smith D.J.;
RT "Cloning of the gbpB gene from Streptococcus mutans.";
RL J. Dent. Res. 79:224-224(2000).
RN [2]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=5J32;
RX MEDLINE=21481971; PubMed=11598068;
RX DOI=10.1128/IAI.69.11.6931-6941.2001;
RA Mattos-Graner R.O., Jin S., King W.F., Chen T., Smith D.J.,
Duncan M.J.;
RT "Cloning of the Streptococcus mutans gene encoding glucan binding
protein B and analysis of genetic diversity and protein production in
clinical isolates.";
RT Infect. Immun. 69:6931-6941(2001).
RL EMBL; AF046410; AAK94500.1; -; Genomic_DNA.
DR InterPro; IPR007921; CHAP.
DR InterPro; IPR009148; SIBA.
DR Pfam; PF05257; CHAP; 1.
DR PRINTS; PR01852; SIBAPROTEIN.
DR PROSITE; PS50311; CHAP; 1.
SQ SEQUENCE 431 AA; 44620 MW; 464FE3B563FB7E51 CRC64;

Query Match 100.0%; Score 90; DB 2; Length 431;
Best Local Similarity 100.0%; Pred. No. 0.00017;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TATEAQPSASSASTAAVAAN 20
|||||
Db 306 TATEAQPSASSASTAAVAAN 325

RESULT 2
Q9AG98_STRMU
ID Q9AG98_STRMU PRELIMINARY; PRT; 431 AA.
AC Q9AG98;
DT 01-JUN-2001 (Tremblrel. 17, Created)
DT 01-JUN-2001 (Tremblrel. 17, Last sequence update)
DT 01-FEB-2005 (Tremblrel. 29, Last annotation update)
DE Immunodominant glycoprotein IDG-60 (Glucan-binding protein B).

GN	Name=gbpB; OrderedLocusNames=SMU.22;
OS	Streptococcus mutans.
OC	Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
OC	Streptococcus.
OC	Streptococcus.
NCBI_TaxID=1309;	
[1]	
RP	NUCLEOTIDE SEQUENCE.
RC	STRAIN=UA159 / ATCC 700610 / Serotype c;
RX	MEDLINE=22295063; PubMed=12397186; DOI=10.1073/pnas.172501299;
RA	Ajidic D.J., McShan W.M., McLaughlin R.E., Savic G., Chang J.,
RA	Carson M.B., Primeaux C., Tian R., Kenton S., Jia H.G., Lin S.P.,
RA	Qian Y., Li S., Zhu H., Najjar F.Z., Lai H., White J., Roe B.A.,
RA	Ferretti J.J.;
RA	"genome sequence of Streptococcus mutans UA159, a cariogenic dental
RT	pathogen.";
RT	Proc. Natl. Acad. Sci. U.S.A. 99:14434-14439 (2002).
RL	EMBL; AE014855; AAN57811.1; -; Genomic_DNA.
DR	InterPro; IPR007921; CHAP.
DR	InterPro; IPR009148; Siba.
DR	Pfam; PF05257; CHAP; 1.
DR	PRINTS; PR01852; SIBAPROTEIN.
DR	PROSITE; PS0911; CHAP; 1.
KW	Complete proteome.
SQ	SEQUENCE 431 AA; 44620 MW; 2D1CA685248CCD3E CRC64;
Query Match 100.0%; Score 90; DB 2; Length 431;	
Best Local Similarity 100.0%; Pred. No. 0.00017;	
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;	
QY	1 TATEAOPSASSASTAAVAAN 20
DB	
DB	306 TATEAOPSASSASTAAVAAN 325
RESULT 4	
QY38V1_STRMU	
ID	QY38V1_STRMU PRELIMINARY; PRT; 432 AA.
AC	QY38V1;
DT	01-DEC-2001 (TrEMBLrel. 19, Created)
DT	01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT	01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE	Glucan-binding protein B.
OS	Streptococcus mutans.
OC	Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
OC	Streptococcus.
NCBI_TaxID=1309;	
[1]	
RP	NUCLEOTIDE SEQUENCE.
RC	STRAIN=3SN1;
RC	Medline=21481971; PubMed=11598068;
RX	MEDLINE=21481971; PubMed=11598068;
RX	DOI=10.1128/IAI.69.11.6931-6941.2001;
RA	Mattos-Graner R.O., Jin S., King W.F., Chen T., Smith D.J.,
RA	Duncan M.J.;
RT	"Cloning of the Streptococcus mutans gene encoding glucan binding
RT	protein B and analysis of the gbpB gene from Streptococcus
RT	clinical isolates.";
RL	Infect. Immun. 69:6931-6941 (2001).
DR	EMBL; AY046413; AAK94503.1; -; Genomic_DNA.
DR	InterPro; IPR007921; CHAP.
DR	InterPro; IPR009148; Siba.
DR	Pfam; PF05257; CHAP; 1.
DR	PRINTS; PR01852; SIBAPROTEIN.
DR	PROSITE; PS0911; CHAP; 1.
SQ	SEQUENCE 432 AA; 44652 MW; 3F88ECB9A1F3BE4F CRC64;
Query Match 93.3%; Score 84; DB 2; Length 432;	
Best Local Similarity 95.0%; Pred. No. 0.0012;	

Matches 19; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 TATAQPSASSASTAAVAAN 20
 ||| |||||||||||||
 Db 307 TATAQPSASSASTAAVAAN 326

RESULT 5

Q938V2_STRMU
 ID Q938V2_STRMU PRELIMINARY; PRT; 432 AA.
 AC Q938V2;
 DT 01-DEC-2001 (TrEMBLrel. 19, Created)
 DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
 DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
 DE Glucan-binding protein B.
 OS Streptococcus mutans.
 OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
 OC Streptococcus.
 OX NCBI_TaxID=1309;
 RN [1]
 RP NUCLEOTIDE SEQUENCE.
 RC STRAIN=15JP2;
 RA Jin S., Duncan M.J., Taubman M.A., Smith D.J.;
 RT "Cloning of the gbpB gene from Streptococcus mutans.";
 RL J. Dent. Res. 79:224-224(2000).
 RN [2]

NUCLEOTIDE SEQUENCE.

RC STRAIN=15JP2;
 RX MEDLINE=21481971; PubMed=11598068;
 RX DOI=10.1128/IAI.69.11.6931-6941.2001;
 RA Mattos-Graner R.O., Jin S., King W.F., Chen T., Smith D.J.,
 RA Duncan M.J.;
 RT "Cloning of the Streptococcus mutans gene encoding glucan binding
 RT protein B and analysis of genetic diversity and protein production in
 RT clinical isolates.";
 RL Infect. Immun. 69:6931-6941(2001).
 DR EMBL; AY046412; AAK94502.1; -; Genomic_DNA.
 DR InterPro; IPR007921; CHAP.
 DR InterPro; IPR009148; Siba.
 DR Pfam; PF05257; CHAP; 1.
 DR PRINTS; PR01852; SIBAPROTEIN.
 DR PROSITE; PS50911; CHAP; 1.
 SQ SEQUENCE 432 AA; 44648 MW; E769B2504AE50E9 CRC64;

Query Match 93.3%; Score 84; DB 2; Length 432;

Best Local Similarity 95.0%; Pred. No. 0.0012;

Matches 19; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 TATAQPSASSASTAAVAAN 20
 ||| |||||||||||||
 Db 307 TATAQPSASSASTAAVAAN 326

RESULT 6

Q938V0_STRMU
 ID Q938V0_STRMU PRELIMINARY; PRT; 431 AA.
 AC Q938V0;
 DT 01-DEC-2001 (TrEMBLrel. 19, Created)
 DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
 DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
 DE Glucan-binding protein B.
 OS Streptococcus mutans.
 OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
 OC Streptococcus.
 OX NCBI_TaxID=1309;
 RN [1]
 RP NUCLEOTIDE SEQUENCE.
 RC STRAIN=SSM3;
 RA Jin S., Duncan M.J., Taubman M.A., Smith D.J.;
 RT "Cloning of the gbpB gene from Streptococcus mutans.";
 RL J. Dent. Res. 79:224-224(2000).
 RN [2]

NUCLEOTIDE SEQUENCE.

RC STRAIN=SSM3;
 RX MEDLINE=21481971; PubMed=11598068;
 RX DOI=10.1128/IAI.69.11.6931-6941.2001;
 RA Mattos-Graner R.O., Jin S., King W.F., Chen T., Smith D.J.,
 RA Duncan M.J.;
 RT "Cloning of the Streptococcus mutans gene encoding glucan binding
 RT protein B and analysis of genetic diversity and protein production in
 RT clinical isolates.";
 RL Infect. Immun. 69:6931-6941(2001).
 DR EMBL; AY046412; AAK94502.1; -; Genomic_DNA.
 DR InterPro; IPR007921; CHAP.
 DR InterPro; IPR009148; Siba.
 DR Pfam; PF05257; CHAP; 1.
 DR PRINTS; PR01852; SIBAPROTEIN.
 DR PROSITE; PS50911; CHAP; 1.
 SQ SEQUENCE 432 AA; 44648 MW; E769B2504AE50E9 CRC64;

RC STRAIN=SSM3;

RX MEDLINE=21481971; PubMed=11598068;
 RX DOI=10.1128/IAI.69.11.6931-6941.2001;
 RA Mattos-Graner R.O., Jin S., King W.F., Chen T., Smith D.J.,
 RA Duncan M.J.;
 RT "Cloning of the Streptococcus mutans gene encoding glucan binding
 RT protein B and analysis of genetic diversity and protein production in
 RT clinical isolates.";
 RL Infect. Immun. 69:6931-6941(2001).
 DR EMBL; AY046412; AAK94504.1; -; Genomic_DNA.
 DR InterPro; IPR007921; CHAP.
 DR InterPro; IPR009148; Siba.
 DR Pfam; PF05257; CHAP; 1.
 DR PRINTS; PR01852; SIBAPROTEIN.
 DR PROSITE; PS50911; CHAP; 1.
 SQ SEQUENCE 431 AA; 44650 MW; 05D38D8D8BC4609F CRC64;

Query Match

Best Local Similarity 90.0%; Pred. No. 0.0023;

Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 TATAQPSASSASTAAVAAN 20
 ||| |||||||||||||
 Db 306 TATAQPSASSASTAAVAAN 325

RESULT 7

Q84JT2_ORYSA
 ID Q84JT2_ORYSA PRELIMINARY; PRT; 166 AA.
 AC Q84JT2;
 DT 01-JUN-2003 (TrEMBLrel. 24, Created)
 DT 01-JUN-2003 (TrEMBLrel. 24, Last sequence update)
 DT 01-FEB-2005 (TrEMBLrel. 29, Last annotation update)
 DE Hypothetical protein OSJNB0096M04.140 (Hypothetical protein
 DE OSJNB0043P23.4).
 GN Name=OSJNB0096M04.140; Synonyms=OSJNB0043P23.4;
 OS Oryza sativa (japonica cultivar-group).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
 OC Ehrhartoideae; Oryzaceae; Oryza.
 OX NCBI_TaxID=39947;
 RN [1]
 RP NUCLEOTIDE SEQUENCE.
 RA Buell C.R., Yuan Q., Ouyang S., Liu J., Gansberger K., Jones K.M.,
 RA Overton II L.L., Tsitrin T., Kim M.M., Bera J.J., Jin S.S.,
 RA Padrosh D.W., Tallon L.J., Koo H., Zismann V., Hsiao J., Blunt S.,
 RA Vanaken S.S., Riedmuller S.B., Utterback T.T., Feldblyum T.V.,
 RA Yang Q.Q., Haas B.J., Suh B.B., Peterson J.J., Quackenbush J.,
 RA White O., Salzberg S.L., Fraser C.M.;
 RL Submitted (JUL-2001) to the EMBL/GenBank/DBJ databases.
 RN [2]

NUCLEOTIDE SEQUENCE.

RA Buell R.;
 RL Submitted (FEB-2003) to the EMBL/GenBank/DBJ databases.
 RN [3]

NUCLEOTIDE SEQUENCE.

RA Buell C.R., Yuan Q., Ouyang S., Liu J., Gansberger K., Jones K.M.,
 RA Overton II L.L., Tsitrin T., Kim M.M., Bera J.J., Jin S.S.,
 RA Padrosh D.W., Tallon L.J., Koo H., Zismann V., Hsiao J., Blunt S.,
 RA Vanaken S.S., Riedmuller S.B., Utterback T.T., Feldblyum T.V.,
 RA Yang Q.Q., Haas B.J., Suh B.B., Peterson J.J., Quackenbush J.,
 RA White O., Salzberg S.L., Fraser C.M.;
 RL Submitted (NOV-2001) to the EMBL/GenBank/DBJ databases.
 RN [4]

NUCLEOTIDE SEQUENCE.

RA Buell R.;
 RL Submitted (MAR-2003) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AC092559; AAO37944.1; -; Genomic DNA.
 DR EMBL; AC099324; AAO73278.1; -; Genomic_DNA.
 DR Granene; Q84JT2;
 DR InterPro; IPR000095; PAKbox/RhoBndng.
 DR Pfam; PF00786; PBD; 1.
 DR PROSITE; PS50108; CRIB; 1.

```

KW Hypothetical protein.
SQ SEQUENCE 166 AA; 17181 MW; 6EC080E8429A16DB CRC64;

Query Match 61.1%; Score 55; DB 2; Length 166;
Best Local Similarity 63.2%; Pred. No. 5.9;
Matches 12; Conservative 3; Mismatches 4; Indels 0; Gaps 0;

QY 1 TATEAQPSSASSASTAAVA 19
Db 35 TTQNDPSSSSASAAAMAA 53

RESULT 8
Q9VKU4 DROME
ID Q9VKU4 DROME PRELIMINARY; PRT; 291 AA.
AC Q9VKU4;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE CG13144-PA.
GN Names=CG13144; ORFNames=CG13144;
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
OX NCBI_TaxID=7227;
[1]
RP NUCLEOTIDE SEQUENCE.
RX MEDLINE=20196006; PubMed=107311132; DOI=10.1126/science.287.5461.2185;
RA Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
RA Amanatides P.G., Scher S.E., Li P.W., Hoskins R.A., Galie R.F.,
RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
RA Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
RA Brandon R.C., Rogers Y.-H.C., Blazej R.G., Champe M., Pfeiffer B.D.,
RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,
RA Abril J.F., Agbayani A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,
RA Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
RA Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,
RA Borkova D., Botchan M.R., Bouck J., Brokstein P., Brothier P.,
RA Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
RA de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
RA Dodson K., Doup L.B., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
RA Durbin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,
RA Foslter C., Gabrielian A.E., Garg N.S., Gelbart W.M., Glasser K.,
RA Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
RA Harris N.L., Harvey D.A., Heiman T.J., Hernandez J.R., Houck J.,
RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,
RA Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
RA Kimmel B., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
RA Lasko P., Lei Y., Levitsky A.A., Li J.H., Li Z., Liang Y., Lin X.,
RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
RA Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,
RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
RA Nelson D.R., Nelson K.A., Nixon K., Nueske D.R., Paclet J.M.,
RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
RA Svirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
RA Wang Z.-Y., Wassarman D.A., Weinstock G.M., Weissbach J.,
RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
RT "The genome sequence of Drosophila melanogaster."
RL Science 287:2185-2195(2000).
[2]
RP NUCLEOTIDE SEQUENCE.
RX MEDLINE=22426065; PubMed=12537568;
RA Celniker S.E., Wheeler D.A., Kronmiller B., Carlson J.W., Halpern A.,
RA Patel S., Adams W., Champe M., Dugan S.P., Frise E., Hodgson A.,
RA George R.A., Hoskins R.A., Laverly T., Muzny D.M., Nelson C.R.,

```

```

RA Pacleb J.M., Park S., Pfeiffer B.D., Richards S., Sodergren E.J.,
RA Svirskas R., Tabor P.E., Wan K., Stapleton M., Sutton G.G., Venter C.,
RA Weinstock G., Scher S.E., Myers E.W., Gibbs R.A., Rubin G.M.;
RT "Finishing a whole-genome shotgun: release 3 of the Drosophila
RT melanogaster euchromatic genome sequence."
RL Genome Biol. 3:RESEARCH0079-RESEARCH0079(2002).
[3]
RN NUCLEOTIDE SEQUENCE.
RP MEDLINE=22426070; PubMed=12537573;
RA Kaminker J.S., Bergman C.M., Kronmiller B., Carlson J.W., Svirskas R.,
RA Patel S., Frise E., Wheeler D.A., Lewis S.E., Rubin G.M.,
RA Ashburner M., Celniker S.E.;
RT "The transposable elements of the Drosophila melanogaster euchromatin:
RT a genomics perspective."
RL Genome Biol. 3:RESEARCH0084.1-RESEARCH0084.20(2002).
[4]
RN NUCLEOTIDE SEQUENCE.
RP MEDLINE=22426069; PubMed=12537572;
RA Misra S., Crosby M.A., Mungall C.J., Matthews B.B., Campbell K.S.,
RA Hradecky P., Huang Y., Kaminker J.S., Millburn G.H., Prochownik S.E.,
RA Smith C.D., Tupy J.L., Whitfield E.J., Bayraktaroglu L., Berman B.P.,
RA Bhattencourt B.R., Celniker S.E., de Grey A.D.N.J., Drysdale R.A.,
RA Harris N.L., Richter J., Russo S., Schroeder A.J., Shu S.Q.,
RA Stapleton M., Yamada C., Ashburner M., Gelbart W.M., Rubin G.M.,
RA Lewis S.E.;
RT "Annotation of the Drosophila melanogaster euchromatic genome: a
RT systematic review."
RL Genome Biol. 3:RESEARCH0083.1-RESEARCH0083.22(2002).
[5]
RN NUCLEOTIDE SEQUENCE.
RP Berkeley Drosophila Genome Project;
RG Celniker S., Carlson J., Wan K., Pfeiffer B., Frise E., George R.,
RA Hoskins R., Stapleton M., Paclet J., Park S., Svirskas R., Smith E.,
RA Yu C., Rubin G.;
RT "Drosophila melanogaster release 4 sequence."
RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.
[6]
RN NUCLEOTIDE SEQUENCE.
RP FlyBase;
RL Submitted (MAR-2005) to the EMBL/GenBank/DBJ databases.
CC -!- INTERACTION:
CC Q9VFW7.CG32473; NBExp=1; IntAct=EBI-149908, EBI-140143;
CC P13496.G1; NBExp=1; IntAct=EBI-149908, EBI-157742;
CC EMBL; AE003628; AAF52965.1; -; Genomic_DNA.
DR IntAct; Q9VKU4; -
DR Ensembl; CG13144; Drosophila melanogaster.
DR FlyBase; FBgn0032260; CG13144.
DR GO; GO:0019028; C: viral capsid; IEA.
DR GO; GO:0005198; F: structural molecule activity; IEA.
DR InterPro; IPR001337; TMV coat. F6C3D3343AF27527 CRC64;
SQ SEQUENCE 291 AA; 33122 MW; F6C3D3343AF27527 CRC64;

Query Match 61.1%; Score 55; DB 2; Length 291;
Best Local Similarity 66.7%; Pred. No. 11;
Matches 12; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

QY 1 TATEAQPSSASSASTAAVA 18
Db 112 TSTSPPTPNASSSSSTAAVA 129

RESULT 9
Q6YXX1 ORYSA
ID Q6YXX1 ORYSA PRELIMINARY; PRT; 700 AA.
AC Q6YXX1;
DT 05-JUL-2004 (TrEMBLrel. 27, Created)
DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE Hypothetical protein OSUNBa0026E05.20.
GN Names=OSUNBa0026E05.20;
OS Oryza sativa (Japanese cultivar-group).
OC Eukaryota; Viridiplantae; Streptophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;

```

Db	678	TAAGSPQASSRSTAARAA	697
RESULT 11			
Q68H99_PIG			
ID	Q68H99_PIG	PRELIMINARY;	PRT; 1229 AA.
AC	Q68H99;		
DT	25-OCT-2004	(TREMBLrel. 28, Created)	
DT	25-OCT-2004	(TREMBLrel. 28, Last sequence update)	
DT	25-OCT-2004	(TREMBLrel. 28, Last annotation update)	
DE	Insulin receptor substrate-1 (Fragment).		
GN	Name=IRS-1;		
OS	Sus scrofa (Pig).		
OC	OS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
OC	Mammalia; Eutheria; Laurasiatheria; Cetartiodactyla; Suina; Suidae;		
OC	Sus.		
OX	NCBI_TaxID=9823;		
OX	NCBI_TaxID=9823;		
RP	[1]		
RP	NUCLEOTIDE SEQUENCE.		
RA	Horn P., Hoj A., Jacobsen M., Vingborg R.K.K., Bendixen C.;		
RA	"Characterization of single nucleotide polymorphisms in the porcine		
RT	insulin receptor substrate-1 (IRS-1) gene."		
RL	Submitted (JUL-2004) to the EMBL/GenBank/DBJ databases.		
CC	-1- SIMILARITY: Contains 1 PH domain.		
DR	ENR; AY690661; AAT99886.1; -; Genomic_DNA.		
DR	SMR; Q68H99; 1-254.		
DR	GO; GO:0005158; P:Insulin receptor binding; IEA.		
DR	GO; GO:0004872; P:Receptor activity; IEA.		
DR	InterPro; IPR002404; Insln_receptorS1.		
DR	InterPro; IPR001849; PH.		
DR	Pfam; PF02174; IRS; 1.		
DR	Pfam; PF00169; PH; 1.		
DR	PRINTS; PR00628; INSULINRSI.		
DR	SMART; SM00233; PH; 1.		
DR	SMART; SM00310; PTBI; 1.		
DR	PROSITE; PPS0003; PH_DOMAIN; 1.		
KW	Receptor.		
FT	NON TER	1	
FT	NON TER	1229	
FT	NON TER	1229 AA;	
FT	NON TER	129798 MW;	
FT	NON TER	8DF1FBC5900B042C	
FT	NON TER	CRC64;	
FT	NON TER	Score 51; DB 2; Length 1229;	
FT	NON TER	Best Local Similarity 66.7%; Pred. No. 1.7e+02;	
FT	NON TER	Matches 12; Conservative 3; Mismatches 3; Indels 0; Gaps	
Qy	2	ATEQPSASSASTAAVAA 19	
Db	1018	ATAAPASATSASSAA	1035
RESULT 12			
Q9U113_LEIMA			
ID	Q9U113_LEIMA	PRELIMINARY;	PRT; 1312 AA.
AC	Q9U113;		
DT	01-MAY-2000	(TREMBLrel. 13, Created)	
DT	01-MAR-2001	(TREMBLrel. 16, Last sequence update)	
DT	01-OCT-2003	(TREMBLrel. 25, Last annotation update)	
DE	Hypothetical protein L4830.01.		
GN	Name=L4830.01;		
OS	Leishmania major.		
OC	Eukaryota; Euglenozoa; Kinetoplastida; Trypanosomatidae; Leishmania		
OC	NCBI_TaxID=5664;		
RP	[1]		
RP	NUCLEOTIDE SEQUENCE.		
RC	STRAIN=Friedlin;		
RC	MEDLINE=98146435; PubMed=9477341;		
RA	Ivens A.C., Lewis S.M., Bagherzadeh A., Zhang L., Chan H.M.,		
RA	Smith D.F.;		
RT	"A physical map of the Leishmania major Friedlin genome.";		
RL	Genome Res. 8:135-145(1998).		
RP	[2]		
RP	NUCLEOTIDE SEQUENCE.		

Best Local Similarity	55.0%;	Pred. No. 3.1e+02;	
Matches	11: Conservative	3: Mismatches	6: Indels
		0: Gaps	0: Gaps


```

RESULT 15
QSLRIO_SILPO
ID QSLRIO_SILPO PRELIMINARY; PRT; 481 AA.
AC QSLRIO;
DT 01-FEB-2005 (TReMBLrel. 29, Created)
DT 01-FEB-2005 (TReMBLrel. 29, Last sequence update)
DT 01-FEB-2005 (TReMBLrel. 29, Last annotation update)
DE Hypothetical protein.
GN OrderedLocusNames=SP02148;
OS Silicibacter pomeroyi.
OC Bacteria; Proteobacteria; Alphaproteobacteria; Rhodobacterales;
OC Rhodobacteraceae; Silicibacter.
OX NCBI_TaxID=89184;
RN [1]
NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
RP STRAIN=DSS-3 / ATCC 700808 / DSM 15171;
RX PubMed=15602564; DOI=10.1038/nature03170;
RA Moran M.A., Buchan A., Gonzalez J.M., Heidelberg J.F., Whitman W.B.,
RA Kiene R.P., Henriksen J.R., King G.M., Belas R., Fuqua C.,
RA Brinkac L.M., Lewis M., Johri S., Weaver B., Pai G., Eisen J.A.,
RA Rahe E., Sheldon W.M., Ye W., Miller T.R., Carlton J., Rasko D.A.,
RA Paulsen I.T., Ren Q., Daugherty S.C., DeBoy R.T., Dodson R.J.,
RA Durkin A.S., Madupu R., Nelson W.C., Sullivan S.A., Rosovitz M.J.,
RA Haft D.H., Selengut J., Ward N.;
RT "Genome sequence of Silicibacter pomeroyi reveals adaptations to the
RT marine environment.";
RL Nature 432:910-913 (2004).
DR EMBL; CP000031; AAV95416.1; -; Genomic DNA.
DR InterPro; IPR006837; DUF610 Y1BQ.
DR Pfam; PF04748; Polysacc_deac_2; 1.
KW Complete proteome; Hypothetical protein.
SQ SEQUENCE 481 AA; 48634 MW; 98388BCA07279D1A CRC64;

Query Match 55.6%; Score 50; DB 2; Length 481;
Best Local Similarity 57.9%; Pred. No. 90;
Matches 11; Conservative 2; Mismatches 6; Indels 0; Gaps 0;

QY 1 TATEAQPSSASSASTAAVAA 19
   |||:|:|:|:|:|
Db 175 TATDAQPAQESASVAEAPA 193

```

Search completed: January 25, 2006, 19:13:35
 Job time : 71.3 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2006 Compugen Ltd.

OM protein - protein search, using sw model

Run on: January 25, 2006, 18:50:05 ; Search time 16.6 Seconds
(without alignments)
99.609 Million cell updates/sec

Title: US-10-797-821-3
Perfect score: 90
Sequence: 1 TATEAQPSSASTAAVAAN 20

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 572060 seqs, 82675679 residues

Total number of hits satisfying chosen parameters: 572060

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued Patents AA:*
1: /cgn2_6/prodata/1/iaa/5_COMB.pep.*
2: /cgn2_6/prodata/1/iaa/6_COMB.pep.*
3: /cgn2_6/prodata/1/iaa/H_COMB.pep.*
4: /cgn2_6/prodata/1/iaa/PC/US COMB.pep.*
5: /cgn2_6/prodata/1/iaa/RE_COMB.pep.*
6: /cgn2_6/prodata/1/iaa/backfile1.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	47	52.2	825	2 US-10-210-428-1	Sequence 1, Appli
2	47	52.2	825	2 US-10-237-551-161	Sequence 161, App
3	47	52.2	826	2 US-09-894-998A-47	Sequence 47, Appl
4	47	52.2	826	2 US-10-237-551-47	Sequence 47, Appl
5	46	51.1	92	2 US-09-344-529-2	Sequence 2, Appli
6	45	50.0	84	2 US-09-252-991A-20152	Sequence 20152, A
7	45	50.0	334	2 US-09-270-767-41733	Sequence 41733, A
8	45	50.0	1024	2 US-09-562-737-50	Sequence 50, Appl
9	44	48.9	158	2 US-09-732-210-858	Sequence 858, App
10	44	48.9	503	2 US-09-487-558B-424	Sequence 424, App
11	43	47.8	98	1 US-07-814-220-2	Sequence 2, Appli
12	43	47.8	98	1 US-07-812-421-2	Sequence 2, Appli
13	43	47.8	107	2 US-09-252-991A-23624	Sequence 23624, A
14	43	47.8	143	2 US-09-252-991A-32580	Sequence 32580, A
15	43	47.8	184	2 US-09-270-767-31979	Sequence 31979, A
16	43	47.8	184	2 US-09-270-767-47196	Sequence 47196, A
17	43	47.8	223	2 US-09-009-816-4	Sequence 4, Appli
18	43	47.8	223	2 US-09-248-796A-24067	Sequence 24067, A
19	43	47.8	365	2 US-09-252-991A-30337	Sequence 30337, A
20	43	47.8	367	2 US-09-009-816-2	Sequence 2, Appli
21	43	47.8	367	2 US-09-557-262-6	Sequence 6, Appli
22	43	47.8	465	2 US-09-270-767-46340	Sequence 46340, A
23	43	47.8	551	2 US-09-252-991A-32760	Sequence 32760, A
24	43	47.8	1024	2 US-09-562-737-43	Sequence 43, Appl
25	43	47.8	4551	2 US-09-320-878-1	Sequence 1, Appli
26	43	47.8	4551	2 US-09-141-908-2	Sequence 2, Appli
27	43	47.8	4551	2 US-09-657-440-1	Sequence 1, Appli

28	43	47.8	4551	2 US-09-793-708-1	Sequence 1, Appli
29	43	47.8	4613	2 US-09-105-537-31	Sequence 31, Appl
30	43	47.8	11877	2 US-09-105-537-6	Sequence 6, Appli
31	42	46.7	110	2 US-09-902-540-12369	Sequence 12369, A
32	42	46.7	141	2 US-09-252-991A-26840	Sequence 26840, A
33	42	46.7	148	2 US-09-252-991A-23009	Sequence 23009, A
34	42	46.7	155	2 US-09-252-991A-25452	Sequence 25452, A
35	42	46.7	187	2 US-09-252-991A-17115	Sequence 17115, A
36	42	46.7	217	2 US-09-270-767-31877	Sequence 31877, A
37	42	46.7	217	2 US-09-270-767-47094	Sequence 47094, A
38	42	46.7	241	2 US-09-640-211A-1039	Sequence 1039, Ap
39	42	46.7	340	4 PCT-US96-02331-13	Sequence 13, Appl
40	42	46.7	540	2 US-09-134-000C-4765	Sequence 4765, Ap
41	42	46.7	594	2 US-09-252-991A-23804	Sequence 23804, A
42	42	46.7	842	4 PCT-US96-02331-15	Sequence 15, Appl
43	42	46.7	851	2 US-09-252-991A-24773	Sequence 24773, A
44	42	46.7	1858	2 US-09-902-540-12643	Sequence 12643, A
45	41.5	46.1	402	1 US-08-477-254A-2	Sequence 2, Appli

ALIGNMENTS

RESULT 1
US-10-210-428-1
; Sequence 1, Application US/10210428
; Patent No. 6814969
; GENERAL INFORMATION:
; APPLICANT: David M. Koelle
; APPLICANT: Nancy A. Hosken
; TITLE OF INVENTION: IMMUNOLOGICALLY SIGNIFICANT HERPES
; FILE REFERENCE: 30967.8-US-U1
; CURRENT APPLICATION NUMBER: US/10/210,428
; CURRENT FILING DATE: 2002-07-31
; PRIOR APPLICATION NUMBER: 60/308,923
; PRIOR FILING DATE: 2001-07-31
; PRIOR APPLICATION NUMBER: 60/309,428
; PRIOR FILING DATE: 2001-08-01
; NUMBER OF SEQ ID NOS: 32
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1
; LENGTH: 825
; TYPE: PRT
; ORGANISM: Herpes simplex virus-2 (HSV-2)
US-10-210-428-1

Query Match 52.2%; Score 47; DB 2; Length 825;
Best Local Similarity 63.2%; Pred. No. 61;
Matches 12; Conservative 2; Mismatches 5; Indels 0; Gaps 0;
QY 1 TATEAQPSSASTAAVA 19
Db 570 SAGAAPPSSASPSSQA 588

RESULT 2
US-10-237-551-161
; Sequence 161, Application US/10237551
; Patent No. 6821519
; GENERAL INFORMATION:
; APPLICANT: Day, Craig H.
; APPLICANT: Hosken, Nancy A.
; APPLICANT: Parsons, Joseph M.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE DIAGNOSIS AND
; FILE REFERENCE: 210121.538C3
; CURRENT APPLICATION NUMBER: US/10/237,551
; CURRENT FILING DATE: 2002-09-06
; NUMBER OF SEQ ID NOS: 254
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 161
; LENGTH: 825

```

; TYPE: PRT
; ORGANISM: HSV2
US-10-237-551-161

Query Match      52.2%; Score 47; DB 2; Length 825;
Best Local Similarity 63.2%; Pred. No. 61;
Matches 12; Conservative 2; Mismatches 5; Indels 0; Gaps 0;

QY 1 TATEAOPSASSASTAAVAA 19
   :|||:||||:|||||
Db 570 SAGAAPSPSSQAQAAVAA 588

RESULT 3
US-09-894-998A-47
; Sequence 47, Application US/09894998A
; Patent No. 6537555
; GENERAL INFORMATION:
; APPLICANT: Hosken, Nancy Ann
; APPLICANT: Craig H. Day
; APPLICANT: Davin C. Dillon
; APPLICANT: McGowan, Patrick
; APPLICANT: Sleath, Paul R.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE DIAGNOSIS AND
; TITLE OF INVENTION: TREATMENT OF HERPES SIMPLEX VIRUS INFECTION
; FILE REFERENCE: 210121.538
; CURRENT APPLICATION NUMBER: US/09/894,998A
; CURRENT FILING DATE: 2001-06-28
; NUMBER OF SEQ ID NOS: 64
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 47
; LENGTH: 826
; TYPE: PRT
; ORGANISM: HSV-2
US-09-894-998A-47

Query Match      52.2%; Score 47; DB 2; Length 826;
Best Local Similarity 63.2%; Pred. No. 61;
Matches 12; Conservative 2; Mismatches 5; Indels 0; Gaps 0;

QY 1 TATEAOPSASSASTAAVAA 19
   :|||:||||:|||||
Db 571 SAGAAPSPSSQAQAAVAA 589

RESULT 4
US-10-237-551-47
; Sequence 47, Application US/10237551
; Patent No. 6821519
; GENERAL INFORMATION:
; APPLICANT: Day, Craig H.
; APPLICANT: Hosken, Nancy A.
; APPLICANT: Parsons, Joseph M.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE DIAGNOSIS AND
; TITLE OF INVENTION: TREATMENT OF HERPES SIMPLEX VIRUS INFECTION
; FILE REFERENCE: 210121.538C3
; CURRENT APPLICATION NUMBER: US/10/237,551
; CURRENT FILING DATE: 2002-09-06
; NUMBER OF SEQ ID NOS: 254
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 47
; LENGTH: 826
; TYPE: PRT
; ORGANISM: HSV-2
US-10-237-551-47

Query Match      52.2%; Score 47; DB 2; Length 826;
Best Local Similarity 63.2%; Pred. No. 61;
Matches 12; Conservative 2; Mismatches 5; Indels 0; Gaps 0;

QY 1 TATEAOPSASSASTAAVAA 19
   :|||:||||:|||||
Db 571 SAGAAPSPSPSSQAQAAVAA 589

RESULT 5
US-09-344-529-2
; Sequence 2, Application US/09344529
; Patent No. 6429293
; GENERAL INFORMATION:
; APPLICANT: Hew, Choy L.
; APPLICANT: HSC Research and Development Limited Partnership
; TITLE OF INVENTION: Sculptin-Type Antifreeze Polypeptides and Nucleic Acids
; FILE REFERENCE: 016252-0026200S
; CURRENT APPLICATION NUMBER: US/09/344,529
; CURRENT FILING DATE: 1999-06-24
; EARLIER APPLICATION NUMBER: US 60/090,794
; EARLIER FILING DATE: 1998-06-26
; EARLIER APPLICATION NUMBER: US 60/095,713
; EARLIER FILING DATE: 1998-08-07
; NUMBER OF SEQ ID NOS: 19
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 2
; LENGTH: 92
; TYPE: PRT
; ORGANISM: Myoxocephalus scorpius
US-09-344-529-2

Query Match      51.1%; Score 46; DB 2; Length 92;
Best Local Similarity 55.6%; Pred. No. 6.7;
Matches 10; Conservative 5; Mismatches 3; Indels 0; Gaps 0;

QY 2 ATEAOPSASSASTAAVAA 19
   |||:||||:|||||
Db 23 ATKAADAASAAAAAIAA 40

RESULT 6
US-09-252-991A-20152
; Sequence 20152, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 20152
; LENGTH: 84
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-20152

Query Match      50.0%; Score 45; DB 2; Length 84;
Best Local Similarity 50.0%; Pred. No. 8.5;
Matches 10; Conservative 5; Mismatches 5; Indels 0; Gaps 0;

QY 1 TATEAOPSASSASTAAVAA 20
   |||:||||:|||||
Db 52 TSTFAVPPADAAASPAVAAS 71

RESULT 7
US-09-270-767-41733
; Sequence 41733, Application US/09270767
; Patent No. 6703491
; GENERAL INFORMATION:
; APPLICANT: Homburger et al.
; TITLE OF INVENTION: Nucleic acids and proteins of Drosophila melanogaster
; FILE REFERENCE: File Reference: 7326-094

```

; CURRENT APPLICATION NUMBER: US/09/270,767
; CURRENT FILING DATE: 1999-03-17
; NUMBER OF SEQ ID NOS: 62517
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 41733
; LENGTH: 334
; TYPE: PRT
; ORGANISM: Drosophila melanogaster
; FEATURE:
; OTHER INFORMATION: Xaa means any amino acid
US-09-270-767-41733

Query Match 50.0%; Score 45; DB 2; Length 334;
Best Local Similarity 55.8%; Pred. No. 42;
Matches 10; Conservative 3; Mismatches 5; Indels 0; Gaps 0;

QY 2 ATEAOPSASSASTAAVAA 19
DB 167 ATAAPTTSNAPTSAANA 184

RESULT 8
US-09-562-737-50
; Sequence 50, Application US/09562737
; Patent No. 6428967
; GENERAL INFORMATION:
; APPLICANT: Herz, Joachim
; APPLICANT: Gotthardt, Michael
; TITLE OF INVENTION: LDL Receptor Signaling Pathways
; FILE REFERENCE: UTSW0708
; CURRENT APPLICATION NUMBER: US/09/562,737
; CURRENT FILING DATE: 2000-05-01
; NUMBER OF SEQ ID NOS: 132
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 50
; LENGTH: 1024
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Synthetic
; OTHER INFORMATION: Sequence
US-09-562-737-50

Query Match 50.0%; Score 45; DB 2; Length 1024;
Best Local Similarity 52.9%; Pred. No. 1.6e+02;
Matches 9; Conservative 4; Mismatches 4; Indels 0; Gaps 0;

QY 3 TEAOPSASSASTAAVAA 19
DB 973 SOAQPSSSAQLALIAA 989

RESULT 9
US-09-732-210-858
; Sequence 858, Application US/09732210
; Patent No. 6573361
; GENERAL INFORMATION:
; APPLICANT: Bunkers, Greg J.
; APPLICANT: Liang, Jihong
; APPLICANT: Mittanck, Cindy A.
; APPLICANT: Seale, Jeffrey W.
; APPLICANT: Wu, Yonnie S.
; TITLE OF INVENTION: Anti-fungal Proteins and Methods for Their Use
; FILE REFERENCE: 38-21(15036)B
; CURRENT APPLICATION NUMBER: US/09/732,210
; CURRENT FILING DATE: 2000-12-07
; PRIOR APPLICATION NUMBER: US 60/169,513
; PRIOR FILING DATE: 1999-12-07
; PRIOR APPLICATION NUMBER: US 60/169,340
; PRIOR FILING DATE: 1999-12-07
; NUMBER OF SEQ ID NOS: 1753
; SEQ ID NO 858
; LENGTH: 158

; TYPE: PRT
; ORGANISM: Puccinia graminis
US-09-732-210-858

Query Match 48.9%; Score 44; DB 2; Length 158;
Best Local Similarity 47.4%; Pred. No. 25;
Matches 9; Conservative 4; Mismatches 6; Indels 0; Gaps 0;

QY 1 TATEAOPSASSASTAAVAA 19
DB 5 SSTKAEPKASSAKTOVAKA 23

RESULT 10
US-09-487-558B-424
; Sequence 424, Application US/09487558B
; Patent No. 6949356
; GENERAL INFORMATION:
; APPLICANT: Busby, Robert
; APPLICANT: Call, Brian
; APPLICANT: Hecht, Peter
; APPLICANT: Holtzman, Doug
; APPLICANT: Madden, Kevin
; APPLICANT: Maxon, Mary
; APPLICANT: Milne, Todd
; APPLICANT: No. 6949356man, Thea
; APPLICANT: Royer, John
; APPLICANT: Salama, Sofie
; APPLICANT: Sherman, Amir
; APPLICANT: Silva, Jeff
; APPLICANT: Summers, Eric
; TITLE OF INVENTION: Methods for Improving Secondary Metabolite Production in Fungi
; FILE REFERENCE: 109272.130
; CURRENT APPLICATION NUMBER: US/09/487,558B
; CURRENT FILING DATE: 2000-01-19
; PRIOR APPLICATION NUMBER: US 60/487,558
; PRIOR FILING DATE: 1999-10-20
; NUMBER OF SEQ ID NOS: 446
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 424
; LENGTH: 503
; TYPE: PRT
; ORGANISM: Saccharomyces cerevisiae
US-09-487-558B-424

Query Match 48.9%; Score 44; DB 2; Length 503;
Best Local Similarity 40.0%; Pred. No. 97;
Matches 8; Conservative 8; Mismatches 4; Indels 0; Gaps 0;

QY 1 TATEAOPSASSASTAAVAA 20
DB 206 TSTSSPSTTSSTGASSSS 225

RESULT 11
US-07-814-220-2
; Sequence 2, Application US/07814220
; Patent No. 5925540
; GENERAL INFORMATION:
; APPLICANT: Caceci, Thomas
; APPLICANT: Toth, Thomas E.
; APPLICANT: Szumanski, Maria B.W.
; TITLE OF INVENTION: SYNTHETIC ANTIFREEZE PEPTIDE AND
; TITLE OF INVENTION: SYNTHETIC GENE CODING FOR ITS PRODUCTION
; NUMBER OF SEQUENCES: 43
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: WHITHAM, CURTIS & WHITHAM
; STREET: Reston Intl. Center, 11800 Sunrise Valley Dr.,
; CITY: Reston
; STATE: VA
; COUNTRY: USA
; ZIP: 20191

```
;
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/814,220
; FILING DATE: 23-DEC-1991
; CLASSIFICATION: 435
; PRIOR APPLICATION NUMBER:
; APPLICATION NUMBER: US 07/588,437
; FILING DATE: 25-SEP-1990
; ATTORNEY/AGENT INFORMATION:
; NAME: Whitham, Michael E.
; REGISTRATION NUMBER: 32,635
; REFERENCE/DOCKET NUMBER: CIT.016
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 703-391-2510
; TELEFAX: 703-391-9035
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 98 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-07-814-220-2

Query Match 47.8%; Score 43; DB 1; Length 98;
Best Local Similarity 55.6%; Pred. No. 20;
Matches 10; Conservative 4; Mismatches 4; Indels 0; Gaps 0;

QY 2 ATEAOPSASSASTAAVAA 19
   ||| :||:|||||
Db 13 ATAATAAAAAATAATAA 30

RESULT 12
US-07-812-421-2
; Sequence 2, Application US/07812421
; Patent No. 5932697
; GENERAL INFORMATION:
; APPLICANT: Caceci, Thomas E.
; APPLICANT: Toch, Thomas E.
; APPLICANT: Szumanski, Maria B.W.
; TITLE OF INVENTION: SYNTHETIC ANTIFREEZE PEPTIDE AND
; TITLE OF INVENTION: SYNTHETIC GENE CODING FOR ITS PRODUCTION
; NUMBER OF SEQUENCES: 43
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: WHITHAM, CURTIS & WHITHAM
; STREET: Reston Intl. Center, 11800 Sunrise Valley Dr.,
; STREET: Suite 900
; CITY: Reston
; STATE: VA
; COUNTRY: USA
; ZIP: 20191
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/812,421
; FILING DATE: 23-DEC-1991
; CLASSIFICATION: 435
; PRIOR APPLICATION NUMBER:
; APPLICATION NUMBER: US 07/588,437
; FILING DATE: 25-SEP-1990
; ATTORNEY/AGENT INFORMATION:
; NAME: Whitham, Michael E.
; REGISTRATION NUMBER: 32,635
; REFERENCE/DOCKET NUMBER: CIT.016
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 703-391-2510
```

```
;
; TELEFAX: 703-391-9035
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 98 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-07-812-421-2

Query Match 47.8%; Score 43; DB 1; Length 98;
Best Local Similarity 55.6%; Pred. No. 20;
Matches 10; Conservative 4; Mismatches 4; Indels 0; Gaps 0;

QY 2 ATEAOPSASSASTAAVAA 19
   ||| :||:|||||
Db 13 ATAATAAAAAATAATAA 30

RESULT 13
US-09-252-991A-23624
; Sequence 23624, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 23624
; LENGTH: 107
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
; US-09-252-991A-23624

Query Match 47.8%; Score 43; DB 2; Length 107;
Best Local Similarity 66.7%; Pred. No. 23;
Matches 10; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

QY 5 AOPSASSASTAAVAA 19
   | :||:|||||
Db 55 ARTSATSSSTAATAA 69

RESULT 14
US-09-252-991A-32580
; Sequence 32580, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 32580
; LENGTH: 143
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
; US-09-252-991A-32580

Query Match 47.8%; Score 43; DB 2; Length 143;
Best Local Similarity 50.0%; Pred. No. 32;
```


THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2006 Compugen Ltd.

OM protein - protein search, using sw model

Run on: January 25, 2006, 19:13:51 ; Search time 60 Seconds
(without alignments)
139.276 Million cell updates/sec

Title: US-10-797-821-3

Perfect score: 90

Sequence: 1 TATEAQPSASSASTAAVAAN 20

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1867569 seqs, 417829326 residues

Total number of hits satisfying chosen parameters: 1867569

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications AA Main:*

- 1: /cgn2_6/prodata1/pubpaa/US07_PUBCOMB.pep:*
- 2: /cgn2_6/prodata1/pubpaa/US08_PUBCOMB.pep:*
- 3: /cgn2_6/prodata1/pubpaa/US09_PUBCOMB.pep:*
- 4: /cgn2_6/prodata1/pubpaa/US10A_PUBCOMB.pep:*
- 5: /cgn2_6/prodata1/pubpaa/US10B_PUBCOMB.pep:*
- 6: /cgn2_6/prodata1/pubpaa/US11_PUBCOMB.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	90	100.0	20	4	US-10-383-930-3
2	90	100.0	20	5	US-10-797-821-3
3	90	100.0	431	4	US-10-383-930-29
4	90	100.0	431	5	US-10-383-930-30
5	90	100.0	431	5	US-10-797-821-29
6	90	100.0	431	5	US-10-797-821-30
7	84	93.3	432	4	US-10-383-930-31
8	84	93.3	432	4	US-10-383-930-32
9	84	93.3	432	5	US-10-797-821-31
10	84	93.3	432	5	US-10-797-821-32
11	82	91.1	431	4	US-10-383-930-33
12	82	91.1	431	5	US-10-797-821-33
13	67	74.4	20	4	US-10-383-930-18
14	67	74.4	20	5	US-10-797-821-18
15	55	61.1	291	6	US-11-097-143-30285
16	54	60.0	1106	4	US-10-282-122A-45050
17	52	57.8	700	4	US-10-437-963-121603
18	50	55.6	213	4	US-10-425-115-331255
19	50	55.6	225	4	US-10-425-115-331259
20	49.5	55.0	339	4	US-10-425-115-262085
21	49.5	55.0	371	4	US-10-425-114-53995
22	49	54.4	189	4	US-10-425-114-54157
23	49	54.4	231	4	US-10-425-114-71651
24	49	54.4	266	4	US-10-425-115-356186
25	49	54.4	607	4	US-10-425-115-269306
26	48	53.3	58	4	US-10-437-963-196822
27	48	53.3	87	4	US-10-425-115-313230

28	48	53.3	416	4	US-10-437-963-167322	Sequence 167322,
29	48	53.3	427	3	US-09-925-297-684	Sequence 684, App
30	48	53.3	446	4	US-10-264-237-1965	Sequence 1965, Ap
31	48	53.3	479	4	US-10-437-963-123073	Sequence 123073,
32	48	53.3	806	4	US-10-437-963-151154	Sequence 151154,
33	47	52.2	109	4	US-10-425-115-359994	Sequence 359994,
34	47	52.2	491	4	US-10-087-192-1182	Sequence 1182, Ap
35	47	52.2	537	5	US-10-723-860-3498	Sequence 3498, Ap
36	47	52.2	644	6	US-11-097-143-38100	Sequence 38100, A
37	47	52.2	686	6	US-11-097-143-14409	Sequence 14409, A
38	47	52.2	825	4	US-10-121-988-161	Sequence 161, App
39	47	52.2	825	4	US-10-200-562-161	Sequence 161, App
40	47	52.2	825	4	US-10-237-551-161	Sequence 161, App
41	47	52.2	825	4	US-10-210-428-1	Sequence 1, Appl
42	47	52.2	825	5	US-10-882-074-1	Sequence 1, Appl
43	47	52.2	825	5	US-10-945-050-161	Sequence 161, App
44	47	52.2	826	3	US-09-894-998-47	Sequence 47, Appl
45	47	52.2	826	4	US-10-121-988-47	Sequence 47, Appl

ALIGNMENTS

RESULT 1

US-10-383-930-3
; Sequence 3, Application US/10383930
; Publication No. US20040127400A1
; GENERAL INFORMATION:
; APPLICANT: Smith, Daniel J
; APPLICANT: Taubman, Martin A
; TITLE OF INVENTION: Immunogenicity of Glucan Binding Protein
; FILE REFERENCE: 25669-018
; CURRENT APPLICATION NUMBER: US/10/383,930
; CURRENT FILING DATE: 2003-03-07
; PRIOR APPLICATION NUMBER: 60/402,483
; PRIOR FILING DATE: 2002-08-08
; PRIOR APPLICATION NUMBER: 60/363,209
; PRIOR FILING DATE: 2002-03-07
; NUMBER OF SEQ ID NOS: 41
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 3
; LENGTH: 20
; TYPE: PRT
; ORGANISM: Streptococcus mutans
US-10-383-930-3

Query Match 100.0%; Score 90; DB 4; Length 20;
Best Local Similarity 100.0%; Pred.No. 4.5e-06; Mismatches 0; Indels 0; Gaps 0;
Matches 20; Conservative 0;
QY 1 TATEAQPSASSASTAAVAAN 20
|||||
Db 1 TATEAQPSASSASTAAVAAN 20
|||||

RESULT 2

US-10-797-821-3
; Sequence 3, Application US/10797821
; Publication No. US20050031633A1
; GENERAL INFORMATION:
; APPLICANT: Smith, Daniel J.
; APPLICANT: Taubman, Martin A.
; TITLE OF INVENTION: Glucan Binding Protein and Glycosyltransferase Immunogens
; FILE REFERENCE: 25669-020
; CURRENT APPLICATION NUMBER: US/10/797,821
; CURRENT FILING DATE: 2004-03-09
; PRIOR APPLICATION NUMBER: 10/383,930
; PRIOR FILING DATE: 2003-03-07
; PRIOR APPLICATION NUMBER: 60/363,209
; PRIOR FILING DATE: 2002-03-07
; PRIOR APPLICATION NUMBER: 60/402,483
; PRIOR FILING DATE: 2002-08-08
; PRIOR APPLICATION NUMBER: 09/290,049

```
; PRIOR FILING DATE: 1999-04-12
; PRIOR APPLICATION NUMBER: 60/081,550
; PRIOR FILING DATE: 1998-04-13
; PRIOR APPLICATION NUMBER: 60/115,142
; PRIOR FILING DATE: 1999-01-08
; NUMBER OF SEQ ID NOS: 45
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 3
; LENGTH: 20
; TYPE: PRT
; ORGANISM: Artificial
; FEATURE:
; OTHER INFORMATION: GpBB peptide
US-10-797-821-3
```

```
Query Match 100.0%; Score 90; DB 5; Length 20;
Best Local Similarity 100.0%; Pred. No. 4.5e-06;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY 1 TATEAQPSASSASTAAVAAN 20
| | | | | | | | | | | | | | | | | |
Db 1 TATEAQPSASSASTAAVAAN 20
```

```
RESULT 3
US-10-383-930-29
; Sequence 29, Application US/10383930
; Publication No. US20040127400A1
; GENERAL INFORMATION:
; APPLICANT: Smith, Daniel J
; APPLICANT: Taubman, Martin A
; TITLE OF INVENTION: Immunogenicity of Glucan Binding Protein
; FILE REFERENCE: 25669-018
; CURRENT APPLICATION NUMBER: US/10/383,930
; CURRENT FILING DATE: 2003-03-07
; PRIOR APPLICATION NUMBER: 60/402,483
; PRIOR FILING DATE: 2002-08-08
; PRIOR APPLICATION NUMBER: 60/363,209
; PRIOR FILING DATE: 2002-03-07
; NUMBER OF SEQ ID NOS: 41
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 29
; LENGTH: 431
; TYPE: PRT
; ORGANISM: Streptococcus mutans
US-10-383-930-29
```

```
Query Match 100.0%; Score 90; DB 4; Length 431;
Best Local Similarity 100.0%; Pred. No. 0.00012;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY 1 TATEAQPSASSASTAAVAAN 20
| | | | | | | | | | | | | | | | | |
Db 306 TATEAQPSASSASTAAVAAN 325
```

```
RESULT 4
US-10-383-930-30
; Sequence 30, Application US/10383930
; Publication No. US20040127400A1
; GENERAL INFORMATION:
; APPLICANT: Smith, Daniel J
; APPLICANT: Taubman, Martin A
; TITLE OF INVENTION: Immunogenicity of Glucan Binding Protein
; FILE REFERENCE: 25669-018
; CURRENT APPLICATION NUMBER: US/10/383,930
; CURRENT FILING DATE: 2003-03-07
; PRIOR APPLICATION NUMBER: 60/402,483
; PRIOR FILING DATE: 2002-08-08
; PRIOR APPLICATION NUMBER: 60/363,209
; PRIOR FILING DATE: 2002-03-07
; NUMBER OF SEQ ID NOS: 41
; SOFTWARE: PatentIn version 3.2
```

```
; SEQ ID NO 30
; LENGTH: 431
; TYPE: PRT
; ORGANISM: Streptococcus mutans
US-10-383-930-30
```

```
Query Match 100.0%; Score 90; DB 4; Length 431;
Best Local Similarity 100.0%; Pred. No. 0.00012;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY 1 TATEAQPSASSASTAAVAAN 20
| | | | | | | | | | | | | | | | | |
Db 306 TATEAQPSASSASTAAVAAN 325
```

```
RESULT 5
US-10-797-821-29
; Sequence 29, Application US/10797821
; Publication No. US20050031633A1
; GENERAL INFORMATION:
; APPLICANT: Smith, Daniel J.
; APPLICANT: Taubman, Martin A.
; TITLE OF INVENTION: Glucan Binding Protein and Glycosyltransferase Immunogens
; FILE REFERENCE: 25669-020
; CURRENT APPLICATION NUMBER: US/10/797,821
; CURRENT FILING DATE: 2004-03-09
; PRIOR APPLICATION NUMBER: 10/383,930
; PRIOR FILING DATE: 2003-03-07
; PRIOR APPLICATION NUMBER: 60/363,209
; PRIOR FILING DATE: 2002-03-07
; PRIOR APPLICATION NUMBER: 60/402,483
; PRIOR FILING DATE: 2002-08-08
; PRIOR APPLICATION NUMBER: 09/290,049
; PRIOR FILING DATE: 1999-04-12
; PRIOR APPLICATION NUMBER: 60/081,550
; PRIOR FILING DATE: 1998-04-13
; PRIOR APPLICATION NUMBER: 60/115,142
; PRIOR FILING DATE: 1999-01-08
; NUMBER OF SEQ ID NOS: 45
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 29
; LENGTH: 431
; TYPE: PRT
; ORGANISM: Streptococcus mutans
US-10-797-821-29
```

```
Query Match 100.0%; Score 90; DB 5; Length 431;
Best Local Similarity 100.0%; Pred. No. 0.00012;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY 1 TATEAQPSASSASTAAVAAN 20
| | | | | | | | | | | | | | | | | |
Db 306 TATEAQPSASSASTAAVAAN 325
```

```
RESULT 6
US-10-797-821-30
; Sequence 30, Application US/10797821
; Publication No. US20050031633A1
; GENERAL INFORMATION:
; APPLICANT: Smith, Daniel J.
; APPLICANT: Taubman, Martin A.
; TITLE OF INVENTION: Glucan Binding Protein and Glycosyltransferase Immunogens
; FILE REFERENCE: 25669-020
; CURRENT APPLICATION NUMBER: US/10/797,821
; CURRENT FILING DATE: 2004-03-09
; PRIOR APPLICATION NUMBER: 10/383,930
; PRIOR FILING DATE: 2003-03-07
; PRIOR APPLICATION NUMBER: 60/363,209
; PRIOR FILING DATE: 2002-03-07
; PRIOR APPLICATION NUMBER: 60/402,483
; PRIOR FILING DATE: 2002-08-08
; PRIOR APPLICATION NUMBER: 09/290,049
```


1 GENERATED INFORMATION:
 2 APPLICANT: Smith, Daniel J.
 3 APPLICANT: Taubman, Martin A.
 4 TITLE OF INVENTION: Glucan Binding Protein and Glycosyltransferase Immunogens
 5
 6 FILE REFERENCE: 25669-020
 7 CURRENT APPLICATION NUMBER: US/10/797,821
 8 CURRENT FILING DATE: 2004-03-09
 9 PRIOR APPLICATION NUMBER: 10/383,930
 10 PRIOR FILING DATE: 2003-03-07
 11 PRIOR APPLICATION NUMBER: 60/363,209
 12 PRIOR FILING DATE: 2002-03-07
 13 PRIOR APPLICATION NUMBER: 60/402,483
 14 PRIOR FILING DATE: 2002-08-08
 15 PRIOR APPLICATION NUMBER: 09/290,049
 16 PRIOR FILING DATE: 1999-04-12
 17 PRIOR APPLICATION NUMBER: 60/081,550
 18 PRIOR FILING DATE: 1998-04-13

```

RESULT 14
US-10-797-821-18
; Sequence 18, Application US/10797821
; Publication No. US20050031633A1
; GENERAL INFORMATION:
; APPLICANT: Smith, Daniel J.
; APPLICANT: Taubman, Martin A.
; TITLE OF INVENTION: Glucan Binding Protein
; FILE REFERENCE: 25669-020
; CURRENT APPLICATION NUMBER: US/10/797,821
; CURRENT FILING DATE: 2004-03-09
; PRIOR APPLICATION NUMBER: 10/383,930
; PRIOR FILING DATE: 2003-03-07
; PRIOR APPLICATION NUMBER: 60/363,209
; PRIOR FILING DATE: 2002-03-07
; PRIOR APPLICATION NUMBER: 60/402,483
; PRIOR FILING DATE: 2002-08-08
; PRIOR APPLICATION NUMBER: 09/290,049
; PRIOR FILING DATE: 1999-04-12
; PRIOR APPLICATION NUMBER: 60/081,550
; PRIOR FILING DATE: 1998-04-13
; PRIOR APPLICATION NUMBER: 60/115,142

```

; PRIOR FILING DATE: 1999-01-08
; NUMBER OF SEQ ID NOS: 45
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 18
; LENGTH: 20
; TYPE: PRT
; ORGANISM: Artificial
; FEATURE:
; OTHER INFORMATION: GbpB peptide
US-10-797-821-18

Query Match 74.4%; Score 67; DB 5; Length 20;
Best Local Similarity 100.0%; Pred. No. 0.0086;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 6 QPSASSASTAAVAAN 20
| | | | | | | | | | | | | | | | | | | | | |
Db 1 QPSASSASTAAVAAN 15

RESULT 15
US-11-097-143-30285
; Sequence 30285, Application US/11097143
; Publication No. US20050208558A1
; GENERAL INFORMATION:
; APPLICANT: Venter, J. Craig
; APPLICANT: et al.
; TITLE OF INVENTION: DETECTION KIT, SUCH AS NUCLEIC ACID
; TITLE OF INVENTION: ARRAYS, FOR DETECTING EXPRESSION OF 10,000 OR MORE
; TITLE OF INVENTION: DROSOPHILA GENES.
; FILE REFERENCE: CL000728
; CURRENT APPLICATION NUMBER: US/11/097,143
; CURRENT FILING DATE: 2005-04-04
; PRIOR APPLICATION NUMBER: 60/157,832
; PRIOR FILING DATE: 1999-10-05
; PRIOR APPLICATION NUMBER: 60/160,191
; PRIOR FILING DATE: 1999-10-19
; PRIOR APPLICATION NUMBER: 60/161,932
; PRIOR FILING DATE: 1999-10-28
; PRIOR APPLICATION NUMBER: 60/164,769
; PRIOR FILING DATE: 1999-11-12
; PRIOR APPLICATION NUMBER: 60/173,383
; PRIOR FILING DATE: 1999-12-28
; PRIOR APPLICATION NUMBER: 60/175,693
; PRIOR FILING DATE: 2000-01-12
; PRIOR APPLICATION NUMBER: 60/184,831
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: 60/191,637
; PRIOR FILING DATE: 2000-03-23
; NUMBER OF SEQ ID NOS: 43008
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 30285
; LENGTH: 291
; TYPE: PRT
; ORGANISM: DROSOPHILA
US-11-097-143-30285

Query Match 61.1%; Score 55; DB 6; Length 291;
Best Local Similarity 66.7%; Pred. No. 8.1;
Matches 12; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

QY 1 TATEAQPSASSASTAAVA 18
| | | | | | | | | | | | | | | | | | | | | |
Db 112 TSTSTPTPNASSSSTAAVA 129

Search completed: January 25, 2006, 20:10:01
Job time : 60 secs

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2006 Compugen Ltd.

OM protein - protein search, using sw model

Run on: January 25, 2006, 19:15:26 ; Search time 5.5 Seconds
(without alignments)
39.378 Million cell updates/sec

Title: US-10-797-821-3

Perfect score: 90

Sequence: 1 TATEAQPSASSASTAAVAAN 20

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 75621 seqs, 10829074 residues

Total number of hits satisfying chosen parameters: 75621

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

Published Applications AA New:*

- 1: /cgn2_6/prodata/2/pubpaa/US08_NEW_PUB.pap:*
- 2: /cgn2_6/prodata/2/pubpaa/US06_NEW_PUB.pap:*
- 3: /cgn2_6/prodata/2/pubpaa/US07_NEW_PUB.pap:*
- 4: /cgn2_6/prodata/2/pubpaa/PCT_NEW_PUB.pap:*
- 5: /cgn2_6/prodata/2/pubpaa/US09_NEW_PUB.pap:*
- 6: /cgn2_6/prodata/2/pubpaa/US10_NEW_PUB.pap:*
- 7: /cgn2_6/prodata/2/pubpaa/US11_NEW_PUB.pap:*
- 8: /cgn2_6/prodata/2/pubpaa/US60_NEW_PUB.pap:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Match	Query Length	DB ID	Description
1	90	100.0	431	7	US-11-052-554A-210
2	48	53.3	441	6	US-10-821-234-1668
3	45	50.0	118	7	US-11-156-084-77
4	45	50.0	139	7	US-11-156-084-88
5	45	50.0	150	7	US-11-156-084-70
6	45	50.0	162	7	US-11-156-084-75
7	45	50.0	168	7	US-11-156-084-72
8	45	50.0	220	7	US-11-156-084-136
9	45	50.0	468	7	US-11-156-084-135
10	43	47.8	367	6	US-10-131-826A-208
11	43	47.8	387	6	US-10-467-657-50
12	43	47.8	387	6	US-10-467-657-4476
13	43	47.8	387	6	US-10-467-657-7150
14	42	46.7	971	7	US-11-052-554A-3
15	41	45.6	406	7	US-11-139-425-6
16	41	45.6	579	6	US-10-821-234-874
17	41	45.6	663	6	US-10-467-945A-1
18	40	44.4	277	7	US-11-120-308-150
19	40	44.4	314	7	US-11-155-288-6
20	40	44.4	531	7	US-11-010-239-32
21	39	43.3	1123	7	US-11-037-243-77
22	39	43.3	1263	6	US-10-485-517-127
23	39	43.3	1886	6	US-10-515-868-8
24	38	42.2	49	7	US-11-071-259-3
25	38	42.2	168	7	US-11-052-554A-320

26	38	42.2	357	6	US-10-821-234-1325	Sequence 1325, Ap
27	38	42.2	383	7	US-11-168-874-10	Sequence 10, Appl
28	38	42.2	412	7	US-11-168-874-12	Sequence 12, Appl
29	38	42.2	477	7	US-11-024-959-261	Sequence 261, App
30	38	42.2	512	6	US-10-517-939-236	Sequence 236, App
31	38	42.2	579	6	US-10-975-798-4	Sequence 4, Appli
32	38	42.2	1076	6	US-10-131-826A-219	Sequence 219, App
33	38	42.2	1694	7	US-11-052-554A-83	Sequence 83, Appl
34	37.5	41.7	291	6	US-10-883-512-103	Sequence 103, App
35	37.5	41.7	580	6	US-10-975-798-3	Sequence 3, Appli
36	37.5	41.7	3689	7	US-11-075-185-4	Sequence 5, Appli
37	37	41.1	54	7	US-11-150-054A-5	Sequence 1108, Ap
38	37	41.1	202	6	US-10-793-626-1108	Sequence 224, App
39	37	41.1	286	7	US-11-033-039-224	Sequence 11, Appl
40	37	41.1	319	6	US-10-606-302-11	Sequence 193, App
41	37	41.1	379	7	US-11-052-554A-193	Sequence 40, Appl
42	37	41.1	399	6	US-10-838-616-40	Sequence 1014, Ap
43	37	41.1	407	6	US-10-995-561-1014	Sequence 80, Appl
44	37	41.1	430	6	US-10-517-939-80	Sequence 2, Appli
45	37	41.1	431	7	US-11-169-013-2	

ALIGNMENTS

RESULT 1

US-11-052-554A-210
; Sequence 210, Application US/11052554A
; Publication No. US20050288866A1

GENERAL INFORMATION:
; APPLICANT: Sachdeva, et al.

; TITLE OF INVENTION: COMPUTATIONAL METHOD FOR IDENTIFYING ADHESIN AND ADHESIN-LIKE
; TITLE OF INVENTION: PROTEINS OF THERAPEUTIC POTENTIAL
; FILE REFERENCE: 30853/40359A

; CURRENT APPLICATION NUMBER: US/11/052,554A
; CURRENT FILING DATE: 2005-02-07

; PRIOR APPLICATION NUMBER: US 60/589,227
; PRIOR FILING DATE: 2004-07-20

; PRIOR APPLICATION NUMBER: IN 173/DEL/2004
; PRIOR FILING DATE: 2004-02-06

; NUMBER OF SEQ ID NOS: 763
; SOFTWARE: PatentIn version 3.3

; SEQ ID NO 210
; LENGTH: 431
; TYPE: PRT

; ORGANISM: Streptococcus mutans UA159
US-11-052-554A-210

Query Match 100.0%; Score 90; DB 7; Length 431;
Best Local Similarity 100.0%; Pred. No. 8.8e-07;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 TATEAQPSASSASTAAVAAN 20

Db 306 TATEAQPSASSASTAAVAAN 325

RESULT 2

US-10-821-234-1668

; Sequence 1668, Application US/10821234
; Publication No. US20050255114A1

GENERAL INFORMATION:
; APPLICANT: Labat, Ivan

; APPLICANT: Stache-Crain, Birgit
; APPLICANT: Andarmani, Susan

; APPLICANT: Tang, Y. Iom
; TITLE OF INVENTION: Methods for Diagnosis and Treatment of Preeclampsia

; FILE REFERENCE: 821A
; CURRENT APPLICATION NUMBER: US/10/821,234

; CURRENT FILING DATE: 2004-04-07
; PRIOR APPLICATION NUMBER: US 60/462,047

; PRIOR FILING DATE: 2003-04-07
; NUMBER OF SEQ ID NOS: 1704

; SOFTWARE: pt_seq_genes version 1.0
; SEQ ID NO 1668
; LENGTH: 441
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-821-234-1668

Query Match 53.3%; Score 48; DB 6; Length 441;
Best Local Similarity 84.6%; Pred. No. 2.1;
Matches 11; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 2 ATEQPSASSAST 14
|||:|||||||
DB 378 ATVAPSSASSAST 390

RESULT 3

US-11-156-084-77
; Sequence 77, Application US/11156084
; Publication No. US20060010515A1
; GENERAL INFORMATION:

; APPLICANT: Monsanto Technology LLC
; TITLE OF INVENTION: Controlled expression of cytokinin biosynthetic genes leads to
; TITLE OF INVENTION: agronomically interesting phenotypes
; FILE REFERENCE: (38-21)
; CURRENT APPLICATION NUMBER: US/11/156,084
; CURRENT FILING DATE: 2005-06-17
; NUMBER OF SEQ ID NOS: 364
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 77
; LENGTH: 118
; TYPE: PRT
; ORGANISM: Zea mays subsp. mays
US-11-156-084-77

Query Match 50.0%; Score 45; DB 7; Length 118;
Best Local Similarity 56.2%; Pred. No. 1.3;
Matches 9; Conservative 4; Mismatches 3; Indels 0; Gaps 0;

QY 5 AOPSASSASTAAVAAN 20
|||||:|:|
DB 2 AHPSAAVSSSTALAAAN 17

RESULT 4

US-11-156-084-88
; Sequence 88, Application US/11156084
; Publication No. US20060010515A1
; GENERAL INFORMATION:

; APPLICANT: Monsanto Technology LLC
; TITLE OF INVENTION: Controlled expression of cytokinin biosynthetic genes leads to
; TITLE OF INVENTION: agronomically interesting phenotypes
; FILE REFERENCE: (38-21)
; CURRENT APPLICATION NUMBER: US/11/156,084
; CURRENT FILING DATE: 2005-06-17
; NUMBER OF SEQ ID NOS: 364
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 88
; LENGTH: 139
; TYPE: PRT
; ORGANISM: Zea mays subsp. mays
US-11-156-084-88

Query Match 50.0%; Score 45; DB 7; Length 139;
Best Local Similarity 56.2%; Pred. No. 1.6;
Matches 9; Conservative 4; Mismatches 3; Indels 0; Gaps 0;

QY 5 AOPSASSASTAAVAAN 20
|||||:|:|
DB 2 AHPSAAVSSSTALAAAN 17

RESULT 5

US-11-156-084-70
; Sequence 70, Application US/11156084
; Publication No. US20060010515A1
; GENERAL INFORMATION:

; APPLICANT: Monsanto Technology LLC
; TITLE OF INVENTION: Controlled expression of cytokinin biosynthetic genes leads to
; TITLE OF INVENTION: agronomically interesting phenotypes
; FILE REFERENCE: (38-21)
; CURRENT APPLICATION NUMBER: US/11/156,084
; CURRENT FILING DATE: 2005-06-17
; NUMBER OF SEQ ID NOS: 364
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 70
; LENGTH: 150
; TYPE: PRT
; ORGANISM: Zea mays subsp. mays
US-11-156-084-70

Query Match 50.0%; Score 45; DB 7; Length 150;
Best Local Similarity 56.2%; Pred. No. 1.8;
Matches 9; Conservative 4; Mismatches 3; Indels 0; Gaps 0;

QY 5 AOPSASSASTAAVAAN 20
|||||:|:|
DB 2 AHPSAAVSSSTALAAAN 17

RESULT 6

US-11-156-084-75
; Sequence 75, Application US/11156084
; Publication No. US20060010515A1
; GENERAL INFORMATION:

; APPLICANT: Monsanto Technology LLC
; TITLE OF INVENTION: Controlled expression of cytokinin biosynthetic genes leads to
; TITLE OF INVENTION: agronomically interesting phenotypes
; FILE REFERENCE: (38-21)
; CURRENT APPLICATION NUMBER: US/11/156,084
; CURRENT FILING DATE: 2005-06-17
; NUMBER OF SEQ ID NOS: 364
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 75
; LENGTH: 162
; TYPE: PRT
; ORGANISM: Zea mays subsp. mays
US-11-156-084-75

Query Match 50.0%; Score 45; DB 7; Length 162;
Best Local Similarity 56.2%; Pred. No. 1.9;
Matches 9; Conservative 4; Mismatches 3; Indels 0; Gaps 0;

QY 5 AOPSASSASTAAVAAN 20
|||||:|:|
DB 2 AHPSAAVSSSTALAAAN 17

RESULT 7

US-11-156-084-72
; Sequence 72, Application US/11156084
; Publication No. US20060010515A1
; GENERAL INFORMATION:

; APPLICANT: Monsanto Technology LLC
; TITLE OF INVENTION: Controlled expression of cytokinin biosynthetic genes leads to
; TITLE OF INVENTION: agronomically interesting phenotypes
; FILE REFERENCE: (38-21)
; CURRENT APPLICATION NUMBER: US/11/156,084
; CURRENT FILING DATE: 2005-06-17
; NUMBER OF SEQ ID NOS: 364
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 72
; LENGTH: 168
; TYPE: PRT
; ORGANISM: Zea mays subsp. mays
; FEATURE:

; NAME/KEY: misc feature
 ; LOCATION: (160)..(160)
 ; OTHER INFORMATION: Xaa can be any naturally occurring amino acid
 ; FEATURE:
 ; NAME/KEY: misc feature
 ; LOCATION: (162)..(162)
 ; OTHER INFORMATION: Xaa can be any naturally occurring amino acid
 US-11-156-084-72

Query Match 50.0%; Score 45; DB 7; Length 168;
 Best Local Similarity 56.2%; Pred. No. 2;
 Matches 9; Conservative 4; Mismatches 3; Indels 0; Gaps 0;

QY 5 AOPSASSASTAAVAAN 20
 DB 2 AHPSAAVSSTALAA 17

RESULT 8
 US-11-156-084-136
 ; Sequence 136, Application US/11156084
 ; Publication No. US20060010515A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Monsanto Technology LLC
 ; TITLE OF INVENTION: Controlled expression of cytokinin biosynthetic genes leads to
 ; FILE REFERENCE: (38-21)
 ; CURRENT APPLICATION NUMBER: US/11/156,084
 ; CURRENT FILING DATE: 2005-06-17
 ; NUMBER OF SEQ ID NOS: 364
 ; SOFTWARE: PatentIn version 3.2
 ; SEQ ID NO 136
 ; LENGTH: 220
 ; TYPE: PRT
 ; ORGANISM: Zea mays
 ; FEATURE:
 ; NAME/KEY: misc feature
 ; LOCATION: (195)..(195)
 ; OTHER INFORMATION: Xaa can be any naturally occurring amino acid
 US-11-156-084-136

Query Match 50.0%; Score 45; DB 7; Length 220;
 Best Local Similarity 56.2%; Pred. No. 2.8;
 Matches 9; Conservative 4; Mismatches 3; Indels 0; Gaps 0;

QY 5 AOPSASSASTAAVAAN 20
 DB 2 AHPSAAVSSTALAA 17

RESULT 9
 US-11-156-084-135
 ; Sequence 135, Application US/11156084
 ; Publication No. US20060010515A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Monsanto Technology LLC
 ; TITLE OF INVENTION: Controlled expression of cytokinin biosynthetic genes leads to
 ; FILE REFERENCE: (38-21)
 ; CURRENT APPLICATION NUMBER: US/11/156,084
 ; CURRENT FILING DATE: 2005-06-17
 ; NUMBER OF SEQ ID NOS: 364
 ; SOFTWARE: PatentIn version 3.2
 ; SEQ ID NO 135
 ; LENGTH: 468
 ; TYPE: PRT
 ; ORGANISM: Zea mays
 US-11-156-084-135

Query Match 50.0%; Score 45; DB 7; Length 468;
 Best Local Similarity 56.2%; Pred. No. 6.5;
 Matches 9; Conservative 4; Mismatches 3; Indels 0; Gaps 0;

QY 5 AOPSASSASTAAVAAN 20
 DB 2 AHPSAAVSSTALAA 17

RESULT 10
 US-10-131-826A-208
 ; Sequence 208, Application US/10131826A
 ; Publication No. US20050245730A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Baker, Kevin P.
 ; APPLICANT: Beresini, Maureen
 ; APPLICANT: DeForge, Laura
 ; APPLICANT: Desnoyers, Luc
 ; APPLICANT: Filvaroff, Ellen
 ; APPLICANT: Gao, Wei-Qiang
 ; APPLICANT: Gerritsen, Mary E.
 ; APPLICANT: Goddard, Audrey
 ; APPLICANT: Godowski, Paul J.
 ; APPLICANT: Gurney, Austin L.
 ; APPLICANT: Sherwood, Steven
 ; APPLICANT: Smith, Victoria
 ; APPLICANT: Stewart, Timothy A.
 ; APPLICANT: Tumas, Daniel
 ; APPLICANT: Watanabe, Colin K
 ; APPLICANT: Wood, William
 ; APPLICANT: Zhang, Zemin
 ; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
 ; FILE REFERENCE: P3330R1C128
 ; CURRENT APPLICATION NUMBER: US/10/131,826A
 ; CURRENT FILING DATE: 2002-04-24
 ; PRIOR APPLICATION NUMBER: 60/049911
 ; PRIOR FILING DATE: 1997-06-18
 ; PRIOR APPLICATION NUMBER: 60/056974
 ; PRIOR FILING DATE: 1997-08-26
 ; PRIOR APPLICATION NUMBER: 60/059113
 ; PRIOR FILING DATE: 1997-09-17
 ; PRIOR APPLICATION NUMBER: 60/059115
 ; PRIOR FILING DATE: 1997-09-17
 ; PRIOR APPLICATION NUMBER: 60/059117
 ; PRIOR FILING DATE: 1997-09-17
 ; PRIOR APPLICATION NUMBER: 60/059122
 ; PRIOR FILING DATE: 1997-09-17
 ; PRIOR APPLICATION NUMBER: 60/059184
 ; PRIOR FILING DATE: 1997-09-17
 ; PRIOR APPLICATION NUMBER: 60/059263
 ; PRIOR FILING DATE: 1997-09-18
 ; PRIOR APPLICATION NUMBER: 60/059352
 ; PRIOR FILING DATE: 1997-09-19
 ; PRIOR APPLICATION NUMBER: 60/059588
 ; PRIOR FILING DATE: 1997-09-19
 ; Remaining Prior Application data removed - See File Wrapper or PALM.
 ; NUMBER OF SEQ ID NOS: 550
 ; SEQ ID NO 208
 ; LENGTH: 367
 ; TYPE: PRT
 ; ORGANISM: Homo Sapien
 US-10-131-826A-208

Query Match 47.8%; Score 43; DB 6; Length 367;
 Best Local Similarity 58.8%; Pred. No. 10;
 Matches 10; Conservative 2; Mismatches 5; Indels 0; Gaps 0;

QY 3 TEAOPSASSASTAAVA 19
 DB 79 TPSEPSAPSAPAAVPA 95

RESULT 11
 US-10-467-657-50
 ; Sequence 50, Application US/10467657
 ; Publication No. US20050260581A1

```
; GENERAL INFORMATION:
; APPLICANT: CHIRON SpA
; APPLICANT: FONTANA Maria Rita
; APPLICANT: PIZZA Mariagrazia
; APPLICANT: MASIGNANI Vega
; APPLICANT: MONACI Elisabetta
; TITLE OF INVENTION: GONOCOCCAL PROTEINS AND NUCLEIC ACIDS
; FILE REFERENCE:
; CURRENT APPLICATION NUMBER: US/10/467,657
; CURRENT FILING DATE: 2003-08-11
; PRIOR APPLICATION NUMBER: GB-0103424.8
; PRIOR FILING DATE: 2001-02-12
; NUMBER OF SEQ ID NOS: 9218
; SOFTWARE: SeqWin99, version 1.04
; SEQ ID NO 50
; LENGTH: 387
; TYPE: PRT
; ORGANISM: Neisseria gonorrhoeae
US-10-467-657-50

Query Match          47.8%; Score 43; DB 6; Length 387;
Best Local Similarity 83.3%; Pred. No. 11;
Matches 10; Conservative 1; Mismatches 1; Indels 1; Gaps 0;

QY      2 ATEAQPSSASSAS 13
Db      46 ASEAVPSASSAS 57

RESULT 12
US-10-467-657-4476
; Sequence 4476, Application US/10467657
; Publication No. US20050260581A1
; GENERAL INFORMATION:
; APPLICANT: CHIRON SpA
; APPLICANT: FONTANA Maria Rita
; APPLICANT: PIZZA Mariagrazia
; APPLICANT: MASIGNANI Vega
; APPLICANT: MONACI Elisabetta
; TITLE OF INVENTION: GONOCOCCAL PROTEINS AND NUCLEIC ACIDS
; FILE REFERENCE:
; CURRENT APPLICATION NUMBER: US/10/467,657
; CURRENT FILING DATE: 2003-08-11
; PRIOR APPLICATION NUMBER: GB-0103424.8
; PRIOR FILING DATE: 2001-02-12
; NUMBER OF SEQ ID NOS: 9218
; SOFTWARE: SeqWin99, version 1.04
; SEQ ID NO 4476
; LENGTH: 387
; TYPE: PRT
; ORGANISM: Neisseria gonorrhoeae
US-10-467-657-4476

Query Match          47.8%; Score 43; DB 6; Length 387;
Best Local Similarity 83.3%; Pred. No. 11;
Matches 10; Conservative 1; Mismatches 1; Indels 1; Gaps 0;

QY      2 ATEAQPSSASSAS 13
Db      46 ASEAVPSASSAS 57

RESULT 13
US-10-467-657-7150
; Sequence 7150, Application US/10467657
; Publication No. US20050260581A1
; GENERAL INFORMATION:
; APPLICANT: CHIRON SpA
; APPLICANT: FONTANA Maria Rita
; APPLICANT: PIZZA Mariagrazia
; APPLICANT: MASIGNANI Vega
; APPLICANT: MONACI Elisabetta
; TITLE OF INVENTION: GONOCOCCAL PROTEINS AND NUCLEIC ACIDS
```

```
; FILE REFERENCE:
; CURRENT APPLICATION NUMBER: US/10/467,657
; CURRENT FILING DATE: 2003-08-11
; PRIOR APPLICATION NUMBER: GB-0103424.8
; PRIOR FILING DATE: 2001-02-12
; NUMBER OF SEQ ID NOS: 9218
; SOFTWARE: SeqWin99, version 1.04
; SEQ ID NO 7150
; LENGTH: 387
; TYPE: PRT
; ORGANISM: Neisseria gonorrhoeae
US-10-467-657-7150

Query Match          47.8%; Score 43; DB 6; Length 387;
Best Local Similarity 83.3%; Pred. No. 11;
Matches 10; Conservative 1; Mismatches 1; Indels 1; Gaps 0;

QY      2 ATEAQPSSASSAS 13
Db      46 ASEAVPSASSAS 57

RESULT 14
US-11-052-554A-3
; Sequence 3, Application US/11052554A
; Publication No. US2005028866A1
; GENERAL INFORMATION:
; APPLICANT: Sachdeva, et al.
; TITLE OF INVENTION: COMPUTATIONAL METHOD FOR IDENTIFYING ADHESIN AND ADHESIN-LIKE
; FILE REFERENCE: 30853/40359A
; CURRENT APPLICATION NUMBER: US/11/052,554A
; CURRENT FILING DATE: 2005-02-07
; PRIOR APPLICATION NUMBER: US 60/589,227
; PRIOR FILING DATE: 2004-07-20
; PRIOR APPLICATION NUMBER: IN 173/DEL/2004
; PRIOR FILING DATE: 2004-02-06
; NUMBER OF SEQ ID NOS: 763
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 3
; LENGTH: 971
; TYPE: PRT
; ORGANISM: Escherichia coli 0157:H7
US-11-052-554A-3

Query Match          46.7%; Score 42; DB 7; Length 971;
Best Local Similarity 58.8%; Pred. No. 43;
Matches 10; Conservative 3; Mismatches 4; Indels 4; Gaps 0;

QY      3 TEAQPSSASSASTAAVAA 19
Db      251 TSAASSASSASSASTAA 267

RESULT 15
US-11-139-425-6
; Sequence 6, Application US/11139425
; Publication No. US20060010518A1
; GENERAL INFORMATION:
; APPLICANT: FELDMANN, Kenneth A
; APPLICANT: NADZAN, Gregory
; APPLICANT: THEISS, Noah
; TITLE OF INVENTION: NUCLEOTIDE SEQUENCES AND POLYPEPTIDES ENCODED THEREBY USEFUL FOR
; FILE REFERENCE: 2750-1597PUS2
; CURRENT APPLICATION NUMBER: US/11/139,425
; CURRENT FILING DATE: 2005-05-27
; PRIOR APPLICATION NUMBER: 60/575,183
; PRIOR FILING DATE: 2004-05-27
; NUMBER OF SEQ ID NOS: 8
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 6
; LENGTH: 406
```

```
; TYPE: PRT
; ORGANISM: Oryza sativa (japonica cultivar-group)
US-11-139-425-6

Query Match      45.6%; Score 41; DB 7; Length 406;
Best Local Similarity 42.1%; Pred. No. 22;
Matches 8; Conservative 7; Mismatches 4; Indels 0; Gaps 0;

QY      1 TATEOPSASSASTAAVAA 19
      |::|::|::|::|::|::|
Db      73 TSSAPSPSASASASAAA 91
```

Search completed: January 25, 2006, 20:11:07
Job time : 5.5 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2006 Compugen Ltd.

OM protein - protein search, using sw model

Run on: January 25, 2006, 18:48:24 ; Search time 68.2 Seconds
(without alignments)
128.850 Million cell updates/sec

Title: US-10-797-821-4
Perfect score: 85
Sequence: 1 LSAVLVSGVTLSSATLSAV 20

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 2443163 seqs, 439378781 residues

Total number of hits satisfying chosen parameters: 2443163

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : A_Geneseq_21.*
1: Geneseq1980s.*
2: Geneseq1990s.*
3: Geneseq2000s.*
4: Geneseq2001s.*
5: Geneseq2002s.*
6: Geneseq2003as.*
7: Geneseq2003bs.*
8: Geneseq2004s.*
9: Geneseq2005s.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	85	100.0	20	7	Add93624 Streptoco
2	85	100.0	20	ADX37247	Adx37247 Streptoco
3	85	100.0	431	7	Add93649 Streptoco
4	85	100.0	431	7	Add93650 Streptoco
5	85	100.0	431	7	Add93653 Streptoco
6	85	100.0	431	9	Adx37272 Streptoco
7	85	100.0	431	9	Adx37273 Streptoco
8	85	100.0	431	9	Adx37276 Streptoco
9	85	100.0	431	9	Aeb91500 Microbial
10	85	100.0	432	7	Add93652 Streptoco
11	85	100.0	432	7	Adx37275 Streptoco
12	84	98.8	432	7	Add93651 Streptoco
13	84	98.8	432	9	Adx37274 Streptoco
14	67	78.8	398	5	Abp25919 Streptoco
15	67	78.8	398	8	Adx83884 S. pyogen
16	67	78.8	398	9	Aeb91542 Microbial
17	53.5	62.9	447	5	Abp29684 Streptoco
18	53.5	62.9	447	8	Adu69524 S agalat
19	53.5	62.9	447	8	Adv88392 Streptoco
20	53.5	62.9	447	8	Adv81808 Streptoco
21	53.5	62.9	447	8	Adv79645 Streptoco
22	48	56.5	177	8	Adx78288 Plant ful
23	48	56.5	456	5	Abv55584 Lactococc
24	47	55.3	27	2	Aaw16500 Lactococc

ALIGNMENTS

RESULT 1									
ADD93624									
ID	ADD93624	standard; peptide; 20 AA.							
AC	ADD93624;								
XX									
DT	29-JAN-2004	(first entry)							
XX									
DE	Streptococcus mutans glucan binding protein-B peptide fragment.								
XX									
KW	Glucan binding protein-B; GbpB; vaccine; anticaries; epitope; immunogen.								
XX									
OS	Streptococcus mutans.								
XX									
PN	WO2003075845-A2.								
XX									
PD	18-SEP-2003.								
XX									
PF	07-MAR-2003; 2003WO-US006962.								
XX									
PR	07-MAR-2002; 2002US-0363209P.								
PR	08-AUG-2002; 2002US-0402483P.								
XX	(FORS-) FORSYTH INST.								
PA	Smith DJ, Taubman MA;								
XX									
PI	WPI; 2003-845091/78.								
DR									
XX									
PT	Composition useful as vaccines for dental caries comprises a fragment of a glucan binding protein-B binding to a major histocompatibility complex class II protein.								
PT									
XX									
PS	Claim 11; Page 10; 49pp; English.								
XX									
CC	The present sequence is that of a peptide comprising amino acid residues 6-25 of the glucan binding protein-B (GbpB) of Streptococcus mutans strain SK32 ADD93649. The peptide binds to a major histocompatibility complex (MHC) class II protein. It was identified as a potential B cell epitope using a matrix-based algorithm for epitope prediction, which was used to search the primary amino acid sequence of GbpB for known MHC class II binding motifs. The peptide can be used in immunogenic compositions and subunit vaccines for dental caries. These comprise MHC class II protein-binding GbpB peptides covalently linked with peptide subunits (preferably from the catalytic domain) of a glucosyltransferase. The compositions are used in a claimed method of eliciting production of an antibody in a mammal. Dieptopic or multieptopic polypeptides can be								

CC prepared synthetically or by recombinant DNA technology. Antibodies
CC raised against MHC class II binding fragments of GbpB can be used in
CC passive immunisation.

SQ Sequence 20 AA;
Query Match 100.0%; Score 85; DB 7; Length 20;
Best Local Similarity 100.0%; Pred. No. 4.7e-06;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 LSAVLVSGVTLSATLSAV 20
| | | | | | | | | | | | | | | | | | | | | |
Db 1 LSAVLVSGVTLSATLSAV 20

RESULT 2
ADX37247
ID ADX37247 standard; peptide; 20 AA.
XX
AC ADX37247;
XX
DT 21-APR-2005 (first entry)
XX
DE Streptococcus mutant glucan binding protein B peptide #4.
XX
KW immunogenicity; immune stimulation; glucan binding protein-B;
KW microparticle; major histocompatibility complex; tooth disease.
XX
OS Streptococcus mutans.
XX
PN US2005031633-A1.
XX
PD 10-FEB-2005.
XX
PF 09-MAR-2004; 2004US-00797821.
XX
PR 13-APR-1998; 98US-0081550P.
PR 08-JAN-1999; 99US-0115142P.
PR 12-APR-1999; 99US-00290049.
PR 07-MAR-2002; 2002US-0363209P.
PR 08-AUG-2002; 2002US-0402483P.
PR 07-MAR-2003; 2003US-00383930.
XX
(SMIT/) SMITH D J.
PA (TAUB/) TAUBMAN M A.
PI Smith DJ, Taubman MA;
XX
XX WPI; 2005-151644/16.
XX
XX New composition comprising a fragment of a glucan binding protein-B
PT (GbpB) that binds to MHC class II protein, and a biocompatible
PT microparticle, useful for producing an antibody (claimed) for immunizing
PT mammals against dental caries.
XX
PS Claim 4; SEQ ID NO 4; 73pp; English.
XX
XX The invention relates to a composition comprising a fragment of a glucan
CC binding protein-B (GbpB) and a biocompatible microparticle, where the
CC fragment binds to a major histocompatibility complex (MHC) class II
CC protein. The composition is useful for producing an antibody for
CC immunizing mammals against dental caries. This sequence corresponds to a
CC Streptococcus mutans GbpB-derived peptide of the invention.
XX
SQ Sequence 20 AA;

Query Match 100.0%; Score 85; DB 9; Length 20;
Best Local Similarity 100.0%; Pred. No. 4.7e-06;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 LSAVLVSGVTLSATLSAV 20
| | | | | | | | | | | | | | | | | | | | | |
Db 1 LSAVLVSGVTLSATLSAV 20

RESULT 3
ADD93649
ID ADD93649 standard; protein; 431 AA.
XX
AC ADD93649;
XX
DT 29-JAN-2004 (first entry)
XX
DE Streptococcus mutans glucan binding protein-B.
XX
KW Glucan binding protein-B; GbpB; vaccine; anticaries; epitope; immunogen.
XX
OS Streptococcus mutans.
XX
FH Key Location/Qualifiers
FT Region 6..25
FT /note= "HLA-binding peptide"
FT Region 16..35
FT /note= "HLA-binding peptide"
FT Region 33..52
FT /note= "HLA-binding peptide"
FT Region 37..56
FT /note= "HLA-binding peptide"
FT Region 48..67
FT /note= "HLA-binding peptide"
FT Region 52..71
FT /note= "HLA-binding peptide"
FT Region 88..107
FT /note= "HLA-binding peptide"
FT Region 113..132
FT /note= "HLA-binding peptide"
FT Region 117..136
FT /note= "HLA-binding peptide"
FT Region 137..156
FT /note= "HLA-binding peptide"
FT Region 174..193
FT /note= "HLA-binding peptide"
FT Region 194..213
FT /note= "HLA-binding peptide"
FT Region 214..233
FT /note= "HLA-binding peptide"
FT Region 248..267
FT /note= "HLA-binding peptide"
FT Region 289..308
FT /note= "HLA-binding peptide"
FT Region 306..325
FT /note= "HLA-binding peptide"
FT Region 311..330
FT /note= "HLA-binding peptide"
FT Region 349..368
FT /note= "HLA-binding peptide"
FT Region 365..384
FT /note= "HLA-binding peptide"
FT Region 383..402
FT /note= "HLA-binding peptide"
FT Region 403..422
FT /note= "HLA-binding peptide"
XX WO2003075845-A2.
XX 18-SEP-2003.
XX 07-MAR-2003; 2003WO-US006962.
XX
PR 07-MAR-2002; 2002US-0363209P.
PR 08-AUG-2002; 2002US-0402483P.
PA (FORS-) FORSYTH INST.
XX Smith DJ, Taubman MA;
XX

WPI; 2003-845091/78.
GENBANK; AY046410.

Composition useful as vaccines for dental caries comprises a fragment of a glucan binding protein-B binding to a major histocompatibility complex class II protein.

Claim 6; Page 7; 49pp; English.

The present sequence is the protein sequence of the glucan binding protein-B (GbpB) of Streptococcus mutans strain SK32. The sequence includes a number of human leucocyte antigen (HLA)-binding epitopes. The invention provides immunogenic compositions and vaccines for dental caries. The compositions comprise major histocompatibility complex (MHC) class II protein-binding GbpB peptides, especially HLA-binding peptides, covalently linked with peptide subunits of a glucosyltransferase. The compositions are used in a claimed method of eliciting production of an antibody in a mammal. Diepitopic or multi-epitopic polypeptides can be prepared synthetically or by recombinant DNA technology. Antibodies raised against MHC class II binding fragments of GbpB can be used in passive immunisation.

Sequence 431 AA;

Query Match 100.0%; Score 85; DB 7; Length 431;
Best Local Similarity 100.0%; Pred. No. 0.00022;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 LSAVLVSGVTLSSTATTLSAV 20
6 LSAVLVSGVTLSSTATTLSAV 25

Db

RESULT 4

ADD93650

ID ADD93650 standard; protein; 431 AA.

AC ADD93650;

XX

29-JAN-2004 (first entry)

Streptococcus mutans glucan binding protein-B.

Glucan binding protein-B; GbpB; vaccine; anticaries; epitope; immunogen.

Streptococcus mutans.

W02003075845-A2.

18-SEP-2003.

07-MAR-2003; 2003WO-US006962.

07-MAR-2002; 2002US-0363209P.

08-AUG-2002; 2002US-0402483P.

(FORS-) FORSYTH INST.

Smith DJ, Taubman MA;

OS Streptococcus mutans.

W02003075845-A2.

18-SEP-2003.

07-MAR-2003; 2003WO-US006962.

07-MAR-2002; 2002US-0363209P.

08-AUG-2002; 2002US-0402483P.

(FORS-) FORSYTH INST.

Smith DJ, Taubman MA;

WPI; 2003-845091/78.
GENBANK; AY046411.

Composition useful as vaccines for dental caries comprises a fragment of a glucan binding protein-B binding to a major histocompatibility complex class II protein.

Claim 5; Page 8; 49pp; English.

The present sequence is the protein sequence of the glucan binding protein-B (GbpB) of Streptococcus mutans strain 3Vr4. The sequence includes a number of human leucocyte antigen (HLA)-binding epitopes. The invention provides immunogenic compositions and vaccines for dental caries. The compositions comprise major histocompatibility complex (MHC) class II protein-binding GbpB peptides, especially HLA-binding peptides, covalently linked with peptide subunits of a glucosyltransferase. The compositions are used in a claimed method of eliciting production of an antibody in a mammal. Diepitopic or multi-epitopic polypeptides can be prepared synthetically or by recombinant DNA technology. Antibodies raised against MHC class II binding fragments of GbpB can be used in passive immunisation.

Sequence 431 AA;

Query Match 100.0%; Score 85; DB 7; Length 431;
Best Local Similarity 100.0%; Pred. No. 0.00022;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 LSAVLVSGVTLSSTATTLSAV 20
6 LSAVLVSGVTLSSTATTLSAV 25

Db

RESULT 5

ADD93653

ID ADD93653 standard; protein; 431 AA.

AC ADD93653;

XX

29-JAN-2004 (first entry)

Streptococcus mutans glucan binding protein-B.

Glucan binding protein-B; GbpB; vaccine; anticaries; epitope; immunogen.

Streptococcus mutans.

W02003075845-A2.

18-SEP-2003.

07-MAR-2003; 2003WO-US006962.

07-MAR-2002; 2002US-0363209P.

08-AUG-2002; 2002US-0402483P.

(FORS-) FORSYTH INST.

Smith DJ, Taubman MA;

WPI; 2003-845091/78.
GENBANK; AY046414.

Composition useful as vaccines for dental caries comprises a fragment of a glucan binding protein-B binding to a major histocompatibility complex class II protein.

Claim 5; Page 8-9; 49pp; English.

The present sequence is the protein sequence of the glucan binding protein-B (GbpB) of Streptococcus mutans strain 5SM3. The sequence includes a number of human leucocyte antigen (HLA)-binding epitopes. The invention provides immunogenic compositions and vaccines for dental caries. The compositions comprise major histocompatibility complex (MHC) class II protein-binding GbpB peptides, especially HLA-binding peptides, covalently linked with peptide subunits of a glucosyltransferase. The compositions are used in a claimed method of eliciting production of an antibody in a mammal. Diepitopic or multi-epitopic polypeptides can be prepared synthetically or by recombinant DNA technology. Antibodies raised against MHC class II binding fragments of GbpB can be used in passive immunisation.

Sequence 431 AA;

Query Match 100.0%; Score 85; DB 7; Length 431;
Best Local Similarity 100.0%; Pred. No. 0.00022;

XX Composition useful as vaccines for dental caries comprises a fragment of
PT a glucan binding protein-B binding to a major histocompatibility complex
PT class II protein.
XX
PS Claim 5; Page 8; 49pp; English.
XX
CC The present sequence is the protein sequence of the glucan binding
CC protein-B (GbpB) of Streptococcus mutans strain 3SNI. The sequence
CC includes a number of human leucocyte antigen (HLA)-binding epitopes. The
CC invention provides immunogenic compositions and vaccines for dental
CC caries. The compositions comprise major histocompatibility complex (MHC)
CC class II protein-binding GbpB peptides, especially HLA-binding peptides,
CC covalently linked with peptide subunits of a glucosyltransferase. The
CC compositions are used in a claimed method of eliciting production of an
CC antibody in a mammal. Dieptopic or multiepitopic polypeptides can be
CC prepared synthetically or by recombinant DNA technology. Antibodies
CC raised against MHC class II binding fragments of GbpB can be used in
CC passive immunisation.
XX
SQ Sequence 432 AA;
Query Match 100.0%; Score 85; DB 7; Length 432;
Best Local Similarity 100.0%; Pred. No. 0.00022;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 LSAVLVSGVTLSSTATLSAV 20
Db 6 LSAVLVSGVTLSSTATLSAV 25
|||||
RESULT 11
ADX37275
ID ADX37275 standard; protein; 432 AA.
XX
AC ADX37275;
XX
DT 21-APR-2005 (first entry)
XX
DE Streptococcus mutant glucan binding protein B variant #4.
XX
KW immunogenicity; immune stimulation; glucan binding protein-B;
KW microparticle; major histocompatibility complex; tooth disease.
XX
OS Streptococcus mutans.
XX
PN US2005031633-A1.
XX
PD 10-FEB-2005.
XX
PF 09-MAR-2004; 2004US-00797821.
XX
PR 13-APR-1998; 98US-0081550P.
PR 08-JAN-1999; 99US-0115142P.
PR 12-APR-1999; 99US-00290049.
PR 07-MAR-2002; 2002US-0363209P.
PR 08-AUG-2002; 2002US-0402483P.
PR 07-MAR-2003; 2003US-00383930.
XX
PA (SMIT/) SMITH D J.
PA (TAUB/) TAUBMAN M A.
XX
PI Smith DJ, Taubman MA;
XX WPI; 2005-151644/16.
XX
XX New composition comprising a fragment of a glucan binding protein-B
PT (GbpB) that binds to MHC class II protein, and a biocompatible
PT microparticle, useful for producing an antibody (claimed) for immunizing
PT mammals against dental caries.
XX
PS Claim 3; SEQ ID NO 32; 73pp; English.
XX

CC The invention relates to a composition comprising a fragment of a glucan
CC binding protein-B (GbpB) and a biocompatible microparticle, where the
CC fragment binds to a major histocompatibility complex (MHC) class II
CC protein. The composition is useful for producing an antibody for
CC immunizing mammals against dental caries. This sequence corresponds to a
CC Streptococcus mutans GbpB protein of the invention.
XX
SQ Sequence 432 AA;
Query Match 100.0%; Score 85; DB 9; Length 432;
Best Local Similarity 100.0%; Pred. No. 0.00022;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 LSAVLVSGVTLSSTATLSAV 20
Db 6 LSAVLVSGVTLSSTATLSAV 25
|||||
RESULT 12
ADD93651
ID ADD93651 standard; protein; 432 AA.
XX
AC ADD93651;
XX
DT 29-JAN-2004 (first entry)
XX
DE Streptococcus mutans glucan binding protein-B.
XX
KW Glucan binding protein-B; GbpB; vaccine; anticaries; epitope; immunogen.
XX
OS Streptococcus mutans.
XX
PN WO2003075845-A2.
XX
PD 18-SEP-2003.
XX
PF 07-MAR-2003; 2003WO-US006962.
XX
PR 07-MAR-2002; 2002US-0363209P.
PR 08-AUG-2002; 2002US-0402483P.
XX
PA (FORS-) FORSYTH INST.
XX
PI Smith DJ, Taubman MA;
XX WPI; 2003-845091/78.
XX GENBANK; AY046412.
XX
PT Composition useful as vaccines for dental caries comprises a fragment of
PT a glucan binding protein-B binding to a major histocompatibility complex
PT class II protein.
XX
XX Claim 5; Page 8; 49pp; English.
XX
CC The present sequence is the protein sequence of the glucan binding
CC protein-B (GbpB) of Streptococcus mutans strain 15JP2. The sequence
CC includes a number of human leucocyte antigen (HLA)-binding epitopes. The
CC invention provides immunogenic compositions and vaccines for dental
CC caries. The compositions comprise major histocompatibility complex (MHC)
CC class II protein-binding GbpB peptides, especially HLA-binding peptides,
CC covalently linked with peptide subunits of a glucosyltransferase. The
CC compositions are used in a claimed method of eliciting production of an
CC antibody in a mammal. Dieptopic or multiepitopic polypeptides can be
CC prepared synthetically or by recombinant DNA technology. Antibodies
CC raised against MHC class II binding fragments of GbpB can be used in
CC passive immunisation.
XX
SQ Sequence 432 AA;
Query Match 98.8%; Score 84; DB 7; Length 432;
Best Local Similarity 95.0%; Pred. No. 0.00031;
Matches 19; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

```

QY 1 LSAVLVSGVTLSAATLSAV 20
Db 6 LSAVLVSGVTLSAATLSAI 25

RESULT 13
AD37274
ID ADX37274 standard; protein; 432 AA.
XX AC ADX37274;
XX DT 21-APR-2005 (first entry)
XX DE Streptococcus mutant glucan binding protein B variant #3.
XX KW immunogenicity; immune stimulation; glucan binding protein-B;
KW microparticle; major histocompatibility complex; tooth disease.
XX OS Streptococcus mutans.
XX PA US2005031633-A1.
XX PN 10-FEB-2005.
XX PD 09-MAR-2004; 2004US-00797821.
XX PF 13-APR-1998; 98US-0081550P.
XX PR 08-JAN-1999; 99US-0115142P.
XX PR 12-APR-1999; 99US-00290049.
XX PR 07-MAR-2002; 2002US-0363209P.
XX PR 08-AUG-2002; 2002US-0402483P.
XX PR 07-MAR-2003; 2003US-00383930.
XX SMITH D J.
XX (TAUBS) TAUBMAN M A.
XX Smith DJ, Taubman MA;
XX WPI; 2005-151644/16.
XX New composition comprising a fragment of a glucan binding protein-B
PT (GbpB) that binds to MHC class II protein, and a biocompatible
PT microparticle, useful for producing an antibody (claimed) for immunizing
PT mammals against dental caries.
XX Claim 3; SEQ ID NO 31; 73pp; English.
XX The invention relates to a composition comprising a fragment of a glucan
CC binding protein-B (GbpB) and a biocompatible microparticle, where the
CC fragment binds to a major histocompatibility complex (MHC) class II
CC protein. The composition is useful for producing an antibody for
CC immunizing mammals against dental caries. This sequence corresponds to a
CC Streptococcus mutans GbpB protein of the invention.
XX Sequence 432 AA;
Query Match 98.8%; Score 84; DB 9; Length 432;
Best Local Similarity 95.0%; Pred. No. 0.00031;
Matches 19; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 LSAVLVSGVTLSAATLSAV 20
Db 6 LSAVLVSGVTLSAATLSAI 25

RESULT 14
ABP25919
ID ABP25919 standard; protein; 398 AA.
XX AC ABP25919;
XX DT 02-JUL-2002 (first entry)
XX KW hyperimmune serum reactive antigen; vaccine; anticaline.

```

```

DE Streptococcus polypeptide SEQ ID NO 1014.
XX Streptococcus; GAS; GBS; group B streptococcus; Streptococcus agalactiae;
KW group A streptococcus; Streptococcus pyogenes; antibacterial;
KW antiinflammatory; infection; vaccine; meningitis; gene therapy.
XX Streptococcus pyogenes.
XX WO200234771-A2.
XX 02-MAY-2002.
XX 29-OCT-2001; 2001WO-GB004789.
XX 27-OCT-2000; 2000GB-00026333.
XX 24-NOV-2000; 2000GB-00028727.
XX 07-MAR-2001; 2001GB-00005640.
XX (CHIR-) CHIRON SPA.
XX (GENO-) INST GENOMIC RES.
XX Telford J, Masighani V, Margarit Y RosI, Grandi G, Fraser C;
XX Tettelin H;
XX WPI; 2002-352536/38.
XX N-PSDB; ABN66550.
XX New Streptococcus protein for the treatment or prevention of infection or
PT disease caused by Streptococcus bacteria, such as meningitis, and for
PT detecting a compound that binds to the protein.
XX Claim 1; Page 3256; 4525pp; English.
XX The invention relates to a protein (ABP25413-ABP30895) from group B
CC streptococcus/GBS (Streptococcus agalactiae) or group A streptococcus/GAS
CC (Streptococcus pyogenes), comprising one of 5483 sequences (SI), given in
CC the specification. The proteins have antibacterial and antiinflammatory
CC activity. (I), nucleic acids encoding (I), ABN66044-ABN71526 and
CC antibodies that bind (I) are used in the manufacture of medicaments for
CC the treatment or prevention of infection or disease caused by
CC Streptococcus bacteria, particularly S. agalactiae and S. pyogenes.
CC Nucleic acids encoding (I) are used to detect Streptococcus in a
CC biological sample. (I) is used to determine whether a compound binds to
CC (I). A composition comprising (I) or a nucleic acid encoding (I), may be
CC used as a vaccine or diagnostic composition. The disease caused by
CC Streptococcus that is prevented or treated may be meningitis. Nucleic
CC acid encoding (I) may be used to recombinantly produce (I) and may be
CC used in gene therapy. Antibodies to (I) are used for affinity
CC chromatography, immunoassays, and distinguishing/identifying
CC Streptococcus proteins
XX Sequence 398 AA;
Query Match 78.8%; Score 67; DB 5; Length 398;
Best Local Similarity 78.9%; Pred. No. 0.11;
Matches 15; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 1 LSAVLVSGVTLSAATLSA 19
Db 6 LSAVLVSGVTLSAATVGA 24

RESULT 15
ADR83884
ID ADR83884 standard; protein; 398 AA.
XX AC ADR83884;
XX DT 02-DEC-2004 (first entry)
XX DE S. pyogenes hyperimmune system reactive antigen SPy0019.
XX KW hyperimmune serum reactive antigen; vaccine; anticaline.

```


GenCore version 5.1.6
Copyright (c) 1993 - 2006 Compugen Ltd.

OM protein - protein search, using sw model

Run on: January 25, 2006, 18:48:55 ; Search time 9.1 Seconds
(without alignments)
211.465 Million cell updates/sec

Title: US-10-797-821-4
Perfect score: 85
Sequence: 1 LSAVLVSGVTLSATLTSVAV 20
Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283416 seqs, 96216763 residues
Total number of hits satisfying chosen parameters: 283416

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : PIR 80.*
1: pir1.*
2: pir2.*
3: pir3.*
4: pir4.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	48	56.5	306	2 AC0258	probable phage pro
2	48	56.5	456	2 E86903	hypothetical prote
3	47	55.3	461	2 JN0097	secreted 45K prote
4	46	54.1	192	2 AG0875	probable membrane
5	44	51.8	601	2 A84256	bacterio-opsin act
6	43	50.6	118	2 E97897	50S ribosomal prot
7	43	50.6	118	2 E95026	ribosomal protein
8	43	50.6	120	2 D69696	ribosomal protein
9	43	50.6	320	2 A40660	outer membrane pro
10	43	50.6	366	2 G97305	spore germination
11	43	50.6	660	2 G97912	hypothetical prote
12	43	50.6	661	2 E95042	hypothetical prote
13	43	50.6	787	2 H71453	hypothetical prote
14	42	49.4	116	2 A87406	ribosomal protein
15	42	49.4	236	2 AB3628	transcription regu
16	42	49.4	325	2 S32308	hypothetical prote
17	42	49.4	363	2 A70636	probable fadB18 pr
18	42	49.4	459	2 F83602	hypothetical prote
19	42	49.4	484	2 A86503	pyruvate kinase [i
20	42	49.4	484	2 A72119	pyruvate kinase CP
21	42	49.4	485	2 S32171	hydroxyneurospore
22	42	49.4	489	2 G64039	hypothetical prote
23	42	49.4	554	2 B82934	hypothetical prote
24	42	49.4	669	2 F82258	DNA ligase VC0971
25	42	49.4	679	2 D64182	DNA ligase (NAD) (
26	42	49.4	737	2 S72442	actin-fragmin kina
27	42	49.4	837	2 G82542	DNA ligase XP2556
28	41	48.2	120	2 AH2813	50S ribosomal prot
29	41	48.2	120	2 D97592	50S ribosomal prot

30	41	48.2	192	2 H91103	probable transport
31	41	48.2	192	2 C85949	hypothetical prote
32	41	48.2	228	2 D64421	hypothetical prote
33	41	48.2	231	2 A87691	conserved hypotet
34	41	48.2	272	2 G75609	hypothetical prote
35	41	48.2	336	2 AH3180	agrobacterium viru
36	41	48.2	369	2 G72730	hypothetical prote
37	41	48.2	427	2 AC0203	chemotaxis MoB pr
38	41	48.2	470	2 D69012	conserved hypotet
39	41	48.2	576	2 AC2195	hypothetical prote
40	41	48.2	586	2 T49441	gamma-glutamyltran
41	41	48.2	598	2 B70036	capsular polysacch
42	41	48.2	746	2 T46821	siderophore recept
43	41	48.2	746	2 A95420	RhA Rhizobactin r
44	41	48.2	1228	2 C98219	proline dehydrogen
45	41	48.2	1228	2 AG3067	proline dehydrogen

ALIGNMENTS

RESULT 1

AC0258
probable phage protein YP02116 [imported] - Yersinia pestis (strain CO92)
C:Species: Yersinia pestis
C:Date: 02-Nov-2001 #sequence_revision 02-Nov-2001 #text_change 09-Jul-2004
C:Accession: AC0258
R:Parkhill, J.; Wren, B.W.; Thomson, N.R.; Titball, R.W.; Holden, M.T.G.; Prentice, M.B.
deno-Tarraga, A.M.; Chillingworth, T.; Cronin, A.; Davies, R.M.; Davis, P.; Dougan, G.;
il, M.; Rutherford, K.; Simmonds, M.; Skelton, J.; Stevens, K.; Whitehead, S.; Barrrell,
Nature 413, 523-527, 2001
A:Title: Genome sequence of Yersinia pestis, the causative agent of plague.
A:Reference number: AB0001; MUID:21470413; PMID:11586360
A:Accession: AC0258
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-306 <KUR>
A:Cross-references: UNIPROT:Q8ZEP4; UNIPARC:UPI00000DCC41; GB:AL590842; PIDN:CAC90927.1
C:Genetics:
A:Gene: YP02116

Query Match 56.5%; Score 48; DB 2; Length 306;
Best Local Similarity 61.1%; Pred. No. 5.9;
Matches 11; Conservative 3; Mismatches 4; Indels 0; Gaps 0;

QY 3 AVLVSQVTLSSATLTSVAV 20
|||:||||: |||: ||
Db 221 AVPTGVTLNKTTLTILAV 238

RESULT 2

E86903
hypothetical protein usp45 [imported] - Lactococcus lactis subsp. lactis (strain IL1403)
C:Species: Lactococcus lactis subsp. lactis
C:Date: 23-Mar-2001 #sequence_revision 23-Mar-2001 #text_change 09-Jul-2004
C:Accession: E86903
R:Boletín, A.; Wincker, P.; Mauger, S.; Jaillon, O.; Malarre, K.; Weissbach, J.; Ehrh
Genome Res. 11, 731-753, 2001
A:Title: The complete genome sequence of the lactic acid bacterium Lactococcus lactis s
A:Reference number: AB6625; MUID:21235186; PMID:11337471
A:Accession: E86903
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-456 <STO>
A:Cross-references: UNIPROT:Q9CDJ1; UNIPARC:UPI00000C6BFC; GB:AE005176; PID:gl27252596;
A:Experimental source: strain IL1403
C:Genetics:
A:Gene: usp45

Query Match 56.5%; Score 48; DB 2; Length 456;
Best Local Similarity 55.0%; Pred. No. 8.9;
Matches 11; Conservative 4; Mismatches 5; Indels 0; Gaps 0;

Best Local Similarity 75.0%; Pred. No. 14;
Matches 9; Conservative 0; Indels 0; Gaps 0;

QY 6 VSGVTLSSATTL 17
I:|||||:
DB 48 VAGVTLASASTL 59

RESULT 8
D69696
ribosomal protein L18 rplR - Bacillus subtilis
C:Species: Bacillus subtilis
C:Date: 05-Dec-1997 #sequence_revision 05-Dec-1997 #text_change 09-Jul-2004
C:Accession: D69696
F:Kunst, F.; Ogasawara, N.; Moszer, I.; Albertini, A.M.; Alloni, G.; Azevedo, V.; Berter
C.; Bron, S.; Broutier, S.; Bruch, C.V.; Caldwell, B.; Capuano, V.; Carter, N.M.; Ch
A.; Ehrlich, S.D.; Emerson, P.T.; Entian, K.D.; Errington, J.; Fabret, C.; Ferrari, E.
Nature 390, 249-256, 1997
A:Authors: Foulger, D.; Fritz, C.; Fujita, Y.; Fuma, S.; Galizzi, A.; Gall
lech, J.; Harwood, C.R.; Henaut, A.; Hilbert, H.; Holsappel, S.; Hosono, S.; Hullo, M.F.
Koetter, P.; Koningstein, G.; Krogh, S.; Kumano, M.; Kurita, K.; Lapidus, A.; Lardinois
A:Authors: Lauber, J.; Lazarevic, V.; Lee, S.M.; Levine, A.; Liu, H.; Masuda, S.; Maue
Y. M.; Ogawa, K.; Ogiwara, A.; Oudega, B.; Park, S.H.; Parro, V.; Pohl, T.M.; Portetel
Rieger, M.; Rivolta, S.; Rocha, E.; Roche, B.; Rose, M.; Sadaie, Y.; Sato, T.; Scanlon
A:Authors: Schleich, S.; Schroeter, R.; Scoffone, F.; Sekiguchi, J.; Sekowska, A.; Ser
akeuchi, M.; Tamakoshi, A.; Tanaka, T.; Terpe, F.; Tognoni, A.; Tosato, V.; Uchiyama
T.; Winters, P.; Wipat, A.; Yamamoto, K.; Yamane, K.; Yasumoto, K.; Yata, K.; Yoshida, A.
A:Authors: Yoshikawa, H.F.; Zumstein, E.; Yoshikawa, H.; Danchin, A.
A:Title: The complete genome sequence of the Gram-positive bacterium Bacillus subtilis.
A:Reference number: A69580; MUID:98044033; PMID:9384377
A:Accession: D69696
A:Status: preliminary; nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-120 <KUN>
A:Cross-references: UNIPROT:P46899; UNIPARC:UPI000005FE2F; GB:Z99104; GB:AL009126; NID:9
A:Experimental source: strain 168
C:Genetics:
A:Gene: rplR
C:Superfamily: Escherichia coli ribosomal protein L18

Query Match 50.6%; Score 43; DB 2; Length 120;
Best Local Similarity 75.0%; Pred. No. 14;
Matches 9; Conservative 0; Indels 0; Gaps 0;

QY 6 VSGVTLSSATTL 17
I:|||||:
DB 48 VNGVTLASASTL 59

RESULT 9
A40660
outer membrane protein OmpL1 - Leptospira alstoni
C:Species: Leptospira alstoni
C:Date: 21-Jan-1994 #sequence_revision 25-Apr-1997 #text_change 20-Sep-1999
C:Accession: A40660
R:Haake, D.A.; Champion, C.I.; Martinich, C.; Shang, E.S.; Bianco, D.R.; Miller, J.N.; I
J. Bacteriol. 175, 4225-4234, 1993
A:Title: Molecular cloning and sequence analysis of the gene encoding OmpL1, a transmem
A:Reference number: A40660; MUID:93308103; PMID:8320237
A:Contents: RMS2
A:Accession: A40660
A:Status: preliminary
A:Molecule type: DNA; protein
A:Residues: 1-320 <HA>
A:Cross-references: UNIPARC:UPI0000057892
A:Note: sequence extracted from NCBI backbone (NCBIN:134517, NCBI:P:134518)
C:Superfamily: Leptospira alstoni outer membrane protein OmpL1

Query Match 50.6%; Score 43; DB 2; Length 320;
Best Local Similarity 58.8%; Pred. No. 37;
Matches 10; Conservative 3; Mismatches 4; Indels 0; Gaps 0;

QY 3 AVLVSQVTLSSATTL 19

A;Authors: Lofus, B.J.; Yang, F.; Smith, H.O.; Venter, J.C.; Dougherty, B.A.; Morrison,
A;Title: Complete Genome Sequence of a virulent isolate of *Streptococcus pneumoniae*.
A;Reference number: A95000; MUID:21357209; PMID:11463916

A;Accession: E95042

A;Status: preliminary

A;Molecule type: DNA

A;Residues: 1-661 <KUR>

A;Cross-references: UNIPARC:UPI0000165A8D; GB:AE005672; PIDN:AAK74534.1; PID:g14971835;

A;Experimental source: strain TIGR4

C;Genetics:

A;Gene: SP0366

Query Match 50.6%; Score 43; DB 2; Length 661;
Best Local Similarity 64.3%; Pred. No. 77;
Matches 9; Conservative 4; Mismatches 1; Indels 0; Gaps 0;

QY 6 VSGVTLSSATTLISA 19

DB 10 LAGVTLAAATTLAA 23

RESULT 13

H71453

hypothetical protein PH0287 - *Pyrococcus horikoshii*

C;Species: *Pyrococcus horikoshii*

C;Date: 14-Aug-1998 #sequence_revision 14-Aug-1998 #text_change 12-Jul-2004

C;Accession: H71453

R;Kawarabayasi, Y.; Sawada, M.; Horikawa, H.; Haikawa, Y.; Hino, Y.; Yamamoto, S.; Sekin

M.; Ohfuku, Y.; Funahashi, T.; Tanaka, T.; Kudoh, Y.; Yamazaki, J.; Kushida, N.; Oguchi

DNA Res. 5, 55-76, 1998

A;Title: Complete sequence and gene organization of the genome of a hyper-thermophilic a

A;Reference number: A71000; MUID:98344137; PMID:9679194

A;Accession: H71453

A;Status: preliminary; nucleic acid sequence not shown; translation not shown

A;Molecule type: DNA

A;Residues: 1-787 <RAW>

A;Cross-references: UNIPROT:O58025; UNIPARC:UPI0000062D86; GB:AP0000001; NID:g3236128; PI

A;Experimental source: strain OT3

A;Note: this accession replaces an interim accession for a sequence replaced by GenBank

C;Genetics:

A;Gene: PH0287

Query Match 50.6%; Score 43; DB 2; Length 787;
Best Local Similarity 50.0%; Pred. No. 92;
Matches 9; Conservative 5; Mismatches 4; Indels 0; Gaps 0;

QY 1 LSAVLVSGVTLSSATTLIS 18

DB 445 LIAVMISGVSLYGATKIT 462

RESULT 14

A87406

ribosomal protein L18 [imported] - *Caulobacter crescentus*

C;Species: *Caulobacter crescentus*

C;Date: 20-Apr-2001 #sequence_revision 20-Apr-2001 #text_change 09-Jul-2004

C;Accession: A87406

R;Nierman, W.C.; Feldblyum, T.V.; Paulsen, I.T.; Nelson, K.E.; Eisen, J.; Heidelberg, J.

B.; Laub, M.T.; DeBoy, R.T.; Dodson, R.J.; Durkin, A.S.; Gwinn, M.L.; Haft, D.H.; Kolon

n, J.; Ermolaeva, M.; White, O.; Salzberg, S.L.; Shapiro, L.; Venter, J.C.; Fraser, C.M.

Proc. Natl. Acad. Sci. U.S.A. 98, 4136-4141, 2001

A;Title: Complete Genome Sequence of *Caulobacter crescentus*.

A;Reference number: A87249; MUID:21173698; PMID:11259647

A;Accession: A87406

A;Status: preliminary

A;Molecule type: DNA

A;Residues: 1-116 <STO>

A;Cross-references: UNIPROT:Q9A8T7; UNIPARC:UPI00000C7318; GB:AE005673; NID:g13422595; F

C;Genetics:

A;Gene: CC1264

C;Superfamily: Escherichia coli ribosomal protein L18

Query Match 49.4%; Score 42; DB 2; Length 116;

Best Local Similarity 75.0%; Pred. No. 19;
Matches 9; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 8 GVTLLSSATTLISA 19

DB 49 GVTLLASASTLLEA 60

RESULT 15

AB3628

transcription regulator, crp family BMEII0947 [imported] - *Brucella melitensis* (strain 1

C;Species: *Brucella melitensis*

C;Date: 01-Feb-2002 #sequence_revision 01-Feb-2002 #text_change 09-Jul-2004

C;Accession: AB3628

R;DeIvecchio, V.G.; Kapatral, V.; Redkar, R.J.; Patra, G.; Mujer, C.; Los, T.; Ivanova, A.

; Mazur, M.; Goltsman, E.; Seikov, E.; Elzer, P.H.; Hagius, S.; O'Callaghan, D.; Letess

Proc. Natl. Acad. Sci. U.S.A. 99, 443-448, 2002

A;Title: The genome sequence of the facultative intracellular pathogen *Brucella melitensis*

A;Reference number: AD3252; PMID:11756688

A;Accession: AB3628

A;Status: preliminary

A;Molecule type: DNA

A;Residues: 1-236 <KUR>

A;Cross-references: UNIPROT:Q8YBF3; UNIPARC:UPI00000586A0; GB:AE008918; PIDN:AAI54189.1;

A;Experimental source: strain LGM

C;Genetics:

A;Gene: BMEII0947

A;Map position: II

Query Match 49.4%; Score 42; DB 2; Length 236;

Best Local Similarity 57.9%; Pred. No. 39;

Matches 11; Conservative 3; Mismatches 5; Indels 0; Gaps 0;

QY 1 LSAVLVSGVTLSSATTLISA 19

DB 83 LAAVLNNDVCLQSAQTLLTA 101

Search completed: January 25, 2006, 19:15:16

Job time : 10.1 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2006 Compugen Ltd.

OM protein - protein search, using sw model

Run on: January 25, 2006, 18:48:24 ; Search time 70.3 Seconds
(without alignments)
200.719 Million cell updates/sec

Title: US-10-797-821-4
Perfect score: 85
Sequence: 1 LSAVLVSGVTLSATLSAV 20

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 2166443 seqs, 705528306 residues

Total number of hits satisfying chosen parameters: 2166443

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Uniprot 05.80.*

1: uniprot_sprot.*

2: uniprot_trembl.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	85	100.0	211	Q9ZAS7_STRMU	Q9zas7 streptococc
2	85	100.0	431	Q938V0_STRMU	Q938v0 streptococc
3	85	100.0	431	Q938V3_STRMU	Q938v3 streptococc
4	85	100.0	431	Q9AG98_STRMU	Q9ag98 streptococc
5	85	100.0	431	Q8DWM3_STRMU	Q8dwm3 streptococc
6	85	100.0	432	Q938V1_STRMU	Q938v1 streptococc
7	84	98.8	432	Q938V2_STRMU	Q938v2 streptococc
8	67	78.8	398	Q5XEL1_STRP6	Q5xel1 streptococc
9	67	78.8	398	Q9ALZ8_STRPY	Q9alz8 streptococc
10	67	78.8	398	Q7CNQ7_STRP8	Q7cnq7 streptococc
11	67	78.8	398	Q8P3L6_STRP3	Q8p3l8 streptococc
12	61	71.8	474	Q5M6K4_STRT2	Q5m6k4 streptococc
13	61	71.8	485	Q5M212_STRT1	Q5m212 streptococc
14	53.5	62.9	447	Q9AKA4_STRAG	Q9akaa4 streptococc
15	53.5	62.9	447	Q8E2H1_STRAS	Q8e2h1 streptococc
16	53.5	62.9	447	Q8E7X9_STRAS	Q8e7x9 streptococc
17	48	56.5	306	Q8ZEP4_YERPE	Q8zep4 yersinia pe
18	48	56.5	456	Q9CDJ1_LACLA	Q9cdj1 lactococcus
19	48	56.5	760	Q5CID7_CRYHO	Q5cid7 cryptospori
20	47	55.3	461	USP45_IACLC	P22865 lactococcus
21	46	54.1	137	Q9AKI1_STRCO	Q9aki1 streptomyce
22	46	54.1	150	Q9R2J3_9ZZZZ	Q9r2j3 plasmid col
23	46	54.1	150	Q07374_9ZZZZ	Q07374 plasmid r64
24	46	54.1	150	Q79VR8_SALTY	Q79vr8 salmonella
25	46	54.1	150	Q6JB00_ECOLI	Q6jb00 escherichia
26	46	54.1	150	Q4FIF9_ENTR	Q4fif9 salmonella
27	46	54.1	192	Q57K42_SALCH	Q57k42 salmonella
28	46	54.1	192	Q5FUI3_SALPA	Q5fui3 salmonella
29	46	54.1	192	Q7CPU4_SALTY	Q7cpu4 salmonella
30	46	54.1	192	Q8XGV9_SALTI	Q8xgv9 salmonella
31	46	54.1	306	Q8LTN3_9CAUD	Q8ltn3 lactococcus

32 46 54.1 341 1 AAPJ_RHILV Q52812 rhizobium 1
33 45 52.9 109 2 Q63Y8_BURPS Q63y8 burkholderi
34 45 52.9 109 2 Q62H95_BURMA Q62h95 burkholderi
35 45 52.9 123 1 RL18_WOLPM Q73ha2 wolbachia p
36 45 52.9 274 2 Q7NCZ1_GLOVI Q7ncz1 gloeobacter
37 45 52.9 313 1 MIRA_PHOIL Q7max8 photorhabdu
38 44 51.8 225 2 Q5TQ37_ANOGA Q5tq37 anopheles g
39 44 51.8 396 2 Q5GSX4_WOLTR Q5gsx4 wolbachia s
40 44 51.8 411 2 Q5FS48_GLOOX Q5fs48 gluconobact
41 44 51.8 413 2 Q84I25_GLOOT Q84i25 clostridium
42 44 51.8 455 2 Q5QWS7_IDILO Q5qws7 idiomarina
43 44 51.8 515 2 Q6TW2_9CAUD Q6iuv2 burkholderi
44 44 51.8 601 2 Q9HQU9_HALSA Q9hqu9 halobacteri
45 44 51.8 660 2 Q5DRK4_XANOR Q5drk4 xanthomonas

ALIGNMENTS

RESULT 1
Q9ZAS7_STRMU
ID Q9ZAS7_STRMU PRELIMINARY; PRT; 211 AA.
AC Q9ZAS7;
DT 01-MAY-1999 (Tremblrel. 10, Created)
DT 01-MAY-1999 (Tremblrel. 10, Last sequence update)
DT 01-OCT-2002 (Tremblrel. 22, Last annotation update)
DE Putative secreted protein (Fragment).
OS Streptococcus mutans.
OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
OC Streptococcus
OX NCBI_TaxID=1309;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=GS-5 Kuramitsu;
RA Peruzzi F., Piggot P.J., Daneo-Moore L.;
RL Submitted (NOV-1996) to the EMBL/GenBank/DBJ databases.
DR EMBL; U78607; AAD00288.1; -; Genomic_DNA.
FT NON_TER 211
SQ SEQUENCE 211 AA; 22803 MW; 4ACE331159CFAFC6 CRC64;

Query Match 100.0%; Score 85; DB 2; Length 211;
Best Local Similarity 100.0%; Pred. No. 4.9e-05; Indels 0; Gaps 0;
Matches 20; Conservative 0; Mismatches 0;

QY 1 LSAVLVSGVTLSATLSAV 20
|||||
DB 6 LSAVLVSGVTLSATLSAV 25

RESULT 2
Q938V0_STRMU
ID Q938V0_STRMU PRELIMINARY; PRT; 431 AA.
AC Q938V0;
DT 01-DEC-2001 (Tremblrel. 19, Created)
DT 01-DEC-2001 (Tremblrel. 19, Last sequence update)
DT 01-MAR-2004 (Tremblrel. 26, Last annotation update)
DE Glucan-binding protein B.
OS Streptococcus mutans.
OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
OC Streptococcus
OX NCBI_TaxID=1309;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=SSM3;
RA Jin S., Duncan M.J., Taubman M.A., Smith D.J.;
RT "Cloning of the gbpB gene from Streptococcus mutans.";
RL J. Dent. Res. 79:224-224 (2000).
RN [2]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=SSM3;
MEDLINE=21481971; PubMed=11598068;
RX DOI=10.1128/JAI.69.11.6931-6941.2001;
RA Mattos-Graner R.O., Jin S., King W.F., Chen T., Smith D.J.,

DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
 DE Putative secreted antigen GbpB/Saga; putative peptidoglycan
 DE hydrolase.
 GN Name=gbpB; OrderedLocusNames=SMU.22;
 OS Streptococcus mutans.
 OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
 OC Streptococcus.
 OX NCBI_TaxID=1309;
 RN [1]
 RP NUCLEOTIDE SEQUENCE.
 RC STRAIN=UA159 / ATCC 700610 / Serotype c;
 RX MEDLINE=22295063; PubMed=12397186; DOI=10.1073/pnas.172501299;
 RA Ajdic D.J., McShan W.M., McLaughlin R.E., Savic G., Chang J.,
 RA Carson M.B., Primeaux C., Tian R., Kenton S., Jia H.G., Lin S.P.,
 RA Qian Y., Li S., Zhu H., Najjar F.Z., Lai H., White J., Roe B.A.,
 RA Ferretti J.J.;
 RT "Genome sequence of Streptococcus mutans UA159, a cariogenic dental
 RT pathogen.";
 RL Proc. Natl. Acad. Sci. U.S.A. 99:14434-14439 (2002).
 DR EMBL; AE014855; AANS7811.1; -; Genomic_DNA.
 DR InterPro; IPR007921; CHAP.
 DR InterPro; IPR009148; Siba.
 DR Pfam; PF05257; CHAP; 1.
 DR PRINTS; PR01852; SIBAPROTEIN.
 DR PROSITE; PS50911; CHAP; 1.
 KW Complete proteome.
 SQ SEQUENCE 431 AA; 44620 MW; 2D1CA685248CCD3E CRC64;

 Query Match 100.0%; Score 85; DB 2; Length 431;
 Best Local Similarity 100.0%; Pred. No. 0.0001;
 Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

 QY 1 LSAVLVSGVTLSSATLSAV 20
 DB 6 LSAVLVSGVTLSSATLSAV 25

 RESULT 6
 Q38V1_STRMU PRELIMINARY; PRT; 432 AA.
 AC Q38V1;
 DT 01-DEC-2001 (TrEMBLrel. 19, Created)
 DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
 DT 01-WAR-2004 (TrEMBLrel. 26, Last annotation update)
 DE Glucan-binding protein B.
 OS Streptococcus mutans.
 OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
 OC Streptococcus.
 OX NCBI_TaxID=1309;
 RN [1]
 RP NUCLEOTIDE SEQUENCE.
 RC STRAIN=3SN1;
 RA Jin S., Duncan M.J., Taubman M.A., Smith D.J.;
 RT "Cloning of the gbpB gene from Streptococcus mutans.";
 RL J. Dent. Res. 79:224-224 (2000).
 RN [2]
 RP NUCLEOTIDE SEQUENCE.
 RC STRAIN=3SN1;
 RX MEDLINE=21481971; PubMed=11598068;
 RX DOI=10.1128/IAI.69.11.6931-6941.2001;
 RA Mattos-Graner R.O., Jin S., King W.F., Chen T., Smith D.J.,
 RA Duncan M.J.;
 RT "Cloning of the Streptococcus mutans gene encoding glucan binding
 RT protein B and analysis of genetic diversity and protein production in
 RT clinical isolates.";
 RL Infect. Immun. 69:6931-6941 (2001).
 DR EMBL; AY046413; AAK94503.1; -; Genomic_DNA.
 DR InterPro; IPR007921; CHAP.
 DR InterPro; IPR009148; Siba.
 DR Pfam; PF05257; CHAP; 1.
 DR PRINTS; PR01852; SIBAPROTEIN.
 DR PROSITE; PS50911; CHAP; 1.
 SQ SEQUENCE 432 AA; 44652 MW; 3F88ECB9A1F3BE4F CRC64;

 Query Match 100.0%; Score 85; DB 2; Length 432;
 Best Local Similarity 100.0%; Pred. No. 0.0001;
 Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

 QY 1 LSAVLVSGVTLSSATLSAV 20
 DB 6 LSAVLVSGVTLSSATLSAV 25

 RESULT 7
 Q38V2_STRMU PRELIMINARY; PRT; 432 AA.
 AC Q38V2;
 DT 01-DEC-2001 (TrEMBLrel. 19, Created)
 DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
 DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
 DE Glucan-binding protein B.
 OS Streptococcus mutans.
 OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
 OC Streptococcus.
 OX NCBI_TaxID=1309;
 RN [1]
 RP NUCLEOTIDE SEQUENCE.
 RC STRAIN=15JP2;
 RA Jin S., Duncan M.J., Taubman M.A., Smith D.J.;
 RT "Cloning of the gbpB gene from Streptococcus mutans.";
 RL J. Dent. Res. 79:224-224 (2000).
 RN [2]
 RP NUCLEOTIDE SEQUENCE.
 RC STRAIN=15JP2;
 RX MEDLINE=21481971; PubMed=11598068;
 RX DOI=10.1128/IAI.69.11.6931-6941.2001;
 RA Mattos-Graner R.O., Jin S., King W.F., Chen T., Smith D.J.,
 RA Duncan M.J.;
 RT "Cloning of the Streptococcus mutans gene encoding glucan binding
 RT protein B and analysis of genetic diversity and protein production in
 RT clinical isolates.";
 RL Infect. Immun. 69:6931-6941 (2001).
 DR EMBL; AY046413; AAK94502.1; -; Genomic_DNA.
 DR InterPro; IPR007921; CHAP.
 DR InterPro; IPR009148; Siba.
 DR Pfam; PF05257; CHAP; 1.
 DR PRINTS; PR01852; SIBAPROTEIN.
 DR PROSITE; PS50911; CHAP; 1.
 SQ SEQUENCE 432 AA; 44648 MW; E769B2504AE50E9 CRC64;

 Query Match 98.8%; Score 84; DB 2; Length 432;
 Best Local Similarity 95.0%; Pred. No. 0.00014;
 Matches 19; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

 QY 1 LSAVLVSGVTLSSATLSAV 20
 DB 6 LSAVLVSGVTLSSATLSAI 25

 RESULT 8
 Q5XEL1_STRP6 PRELIMINARY; PRT; 398 AA.
 AC Q5XEL1;
 DT 25-OCT-2004 (TrEMBLrel. 28, Created)
 DT 25-OCT-2004 (TrEMBLrel. 28, Last sequence update)
 DT 25-OCT-2004 (TrEMBLrel. 28, Last annotation update)
 DE Secreted protein.
 GN OrderedLocusNames=M6_Spy0017;
 OS Streptococcus pyogenes (serotype M6).
 OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
 OC Streptococcus.
 OX NCBI_TaxID=301450;
 RN [1]
 RP NUCLEOTIDE SEQUENCE.
 RC STRAIN=MGAS10394;
 RX PubMed=15272401; DOI=10.1086/422697;

```

RA Banks D.J., Porcella S.F., Barbican K.D., Beres S.B., Phillips L.B.,
RA Voyich J.M., Deleo F.R., Martin J.M., Somerville G.A., Musser J.M.;
RT "Progress toward characterization of the group A Streptococcus
RT metagenome: complete genome sequence of a macrolide-resistant serotype
RT M6 strain.";
RL J. Infect. Dis. 190:727-738(2004).
DR EMBL; CP000003; AAT86152.1; -; Genomic_DNA.
DR InterPro; IPR007921; CHAP.
DR InterPro; IPR009148; Siba.
DR Pfam; PF05257; CHAP; 1.
DR PRINTS; PR01852; SIBAPROTEIN.
DR PROSITE; PS50911; CHAP; 1.
KW Complete proteome.
SQ SEQUENCE 398 AA; 42028 MW; 5DCEDA78CB863B60 CRC64;

Query Match 78.8%; Score 67; DB 2; Length 398;
Best Local Similarity 78.9%; Pred. No. 0.062;
Matches 15; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 1 LSALVSGVTLSSATLSA 19
Db 6 LSALVSGVTLGAAATVGA 24
|||||
[1]

RESULT 9
Q9A1Z8 STRP3 PRELIMINARY; PRT; 398 AA.
AC Q9A1Z8; Q7BH59;
DT 01-JUN-2001 (TrEMBLrel. 17, Created)
DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
DT 10-MAY-2005 (TrEMBLrel. 30, Last annotation update)
DE Secreted protein Siba precursor.
GN OrderedLocusNames=SPY0019;
OS Streptococcus pyogenes.
OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
OC Streptococcus.
OC NCBI_TaxID=1314;
RN NUCLEOTIDE SEQUENCE.
RP STRAIN=SF370 / ATCC 700294 / Serotype M1;
RC MEDLINE=21192684; PubMed=11296296; DOI=10.1073/pnas.071559398;
RX Ferratti J.J., Mcshan W.M., Ajdic D.J., Savic D.J., Savic G., Lyon K.,
RA Primeaux C., Sezate S., Suvorov A.N., Kenton S., Lai H.S., Lin S.P.,
RA Qian Y., Jia H.G., Najjar F.Z., Ren Q., Zhu H., Song L., White J.,
RA Yuan X., Clifton S.W., Roe B.A., McLaughlin R.E.;
RT "Complete genome sequence of an M1 strain of Streptococcus pyogenes.";
RL Proc. Natl. Acad. Sci. U.S.A. 98:4658-4663(2001).
RN [2]
RP NUCLEOTIDE SEQUENCE.
RA Fagan P.K., Reinscheld D., Gottschalk B., Chhatwal G.S.;
RT "Identification and characterization of a novel secreted protein from
RT group A streptococcus.";
RL Submitted (NOV-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; A8006474; AAK33158.1; -; Genomic_DNA.
DR EMBL; AF319999; AAL73135.1; -; Genomic_DNA.
DR InterPro; IPR007921; CHAP.
DR InterPro; IPR009148; Siba.
DR Pfam; PF05257; CHAP; 1.
DR PRINTS; PR01852; SIBAPROTEIN.
DR PROSITE; PS50911; CHAP; 1.
KW Complete proteome; Signal.
FT SIGNAL 1 23 Potential.
SQ SEQUENCE 398 AA; 41899 MW; 28A9B3F7195E969B CRC64;

Query Match 78.8%; Score 67; DB 2; Length 398;
Best Local Similarity 78.9%; Pred. No. 0.062;
Matches 15; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 1 LSALVSGVTLSSATLSA 19
Db 6 LSALVSGVTLGAAATVGA 24
|||||
[1]

```

```

RESULT 10
Q7CNQ7 STRP8 PRELIMINARY; PRT; 398 AA.
AC Q7CNQ7;
DT 05-JUL-2004 (TrEMBLrel. 27, Created)
DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE Putative secreted protein.
DE OrderedLocusNames=spyM18_0020;
OS Streptococcus pyogenes (serotype M18).
OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
OC Streptococcus.
OC NCBI_TaxID=301451;
RN NUCLEOTIDE SEQUENCE.
RP STRAIN=MGAS8232 / Serotype M18;
RX MEDLINE=21927593; PubMed=11917108; DOI=10.1073/pnas.062526099;
RA Smoot J.C., Barbican K.D., Van Gompel J.J., Smoot L.M., Chaussee M.S.,
RA Sylva G.L., Sturdevant D.E., Ricklefs S.M., Porcella S.F.,
RA Parkins L.D., Beres S.B., Campbell D.S., Smith T.M., Zhang Q.,
RA Kapur V., Daly J.A., Veasy L.G., Musser J.M.;
RT "Genome sequence and comparative microarray analysis of serotype M18
RT group A Streptococcus strains associated with acute rheumatic fever
RT outbreaks.";
RL Proc. Natl. Acad. Sci. U.S.A. 99:4668-4673(2002).
DR EMBL; AE009955; AAL96849.1; -; Genomic_DNA.
DR InterPro; IPR007921; CHAP.
DR InterPro; IPR009148; Siba.
DR Pfam; PF05257; CHAP; 1.
DR PRINTS; PR01852; SIBAPROTEIN.
DR PROSITE; PS50911; CHAP; 1.
KW Complete proteome.
SQ SEQUENCE 398 AA; 42028 MW; 5DCEDA78CB863B60 CRC64;

Query Match 78.8%; Score 67; DB 2; Length 398;
Best Local Similarity 78.9%; Pred. No. 0.062;
Matches 15; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 1 LSALVSGVTLSSATLSA 19
Db 6 LSALVSGVTLGAAATVGA 24
|||||
[1]

RESULT 11
Q8P318 STRP3 PRELIMINARY; PRT; 398 AA.
AC Q8P318; Q7CFL7;
DT 01-OCT-2002 (TrEMBLrel. 22, Created)
DT 01-OCT-2002 (TrEMBLrel. 22, Last sequence update)
DT 13-SEP-2005 (TrEMBLrel. 31, Last annotation update)
DE Putative secreted protein.
GN OrderedLocusNames=SP60015, SpyM3_0014;
OS Streptococcus pyogenes (serotype M3).
OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
OC Streptococcus.
OC NCBI_TaxID=301448;
RN NUCLEOTIDE SEQUENCE.
RP STRAIN=SSI-1 / Serotype M3;
RX MEDLINE=22683278; PubMed=12799345; DOI=10.1101/gr.1096703;
RA Nakagawa I., Kurokawa K., Yamashita A., Nakata M., Tomiyasu Y.,
RA Okanashi N., Kawabata S., Yamazaki K., Shiba T., Yasunaga T.,
RA Hayashi H., Hattori M., Hamada S.;
RT "Genome sequence of an M3 strain of Streptococcus pyogenes reveals a
RT large-scale genomic rearrangement in invasive strains and new insights
RT into phage evolution.";
RL Genome Res. 13:1042-1055(2003).
RN [2]
RP NUCLEOTIDE SEQUENCE.
RP STRAIN=MGAS315 / Serotype M3;
RX MEDLINE=22133808; PubMed=12122206; DOI=10.1073/pnas.15298499;
RA Beres S.B., Sylva G.L., Barbican K.D., Lei B., Hoff J.S.,
RA Mammarella N.D., Liu M.-Y., Smoot J.C., Porcella S.F., Parkins L.D.,

```

```

RA Campbell D.S., Smith T.M., McCormick J.K., Leung D.Y.M.,
RA Schlievert P.M., Musser J.M.;
RT "Genome sequence of a serotype M3 strain of group A Streptococcus:
RT phage-encoded toxins, the high-virulence phenotype, and clone
RT emergence.";
RL Proc. Natl. Acad. Sci. U.S.A. 99:10078-10083(2002).
DR EMBL; BA000034; BAC63110.1; -; Genomic_DNA.
DR ENBL; AE014136; AM78621.1; -; Genomic_DNA.
DR InterPro; IPR007921; CHAP.
DR InterPro; IPR009148; Siba.
DR Pfam; PF05257; CHAP; 1.
DR PRINTS; PR01852; SIBAPROTEIN.
DR PROSITE; PS50911; CHAP; 1.
KW Complete proteome.
SQ SEQUENCE 398 AA; 42028 MW; 5DCEDA78CB863B60 CRC64;

Query Match 78.8%; Score 67; DB 2; Length 398;
Best Local Similarity 78.9%; Pred. No. 0.062;
Matches 15; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 1 LSAVLVSGVTLSSAATLSA 19
||| ||||| ||| :|||
DB 6 LSAVLVSGVTLGAATTGVA 24

RESULT 12
Q5M6K4 STRT2
ID Q5M6K4_STRT2 PRELIMINARY; PRT; 474 AA.
AC Q5M6K4;
DT 01-FEB-2005 (TrEMBLrel. 29, Created)
DT 01-FEB-2005 (TrEMBLrel. 29, Last sequence update)
DT 13-SEP-2005 (TrEMBLrel. 31, Last annotation update)
DE Glucan binding protein (PcsB).
GN Name=pcsb; OrderedLocusNames=stu0022;
OS Streptococcus thermophilus (strain ATCC BAA-250 / LMG 18311).
OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
OC Streptococcus.
OX NCBI_TaxID=264199;
RN [1]
RP NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
RX PubMed=15543133; DOI=10.1038/nbt1034;
RA Bolotin A., Quinkis B., Renault P., Sorokin A., Ehrlich S.D.,
RA Kulaukas S., Lapidus A., Goldsman E., Mazur M., Pusch G.D.,
RA Fonstein M., Overbeek R., Kyprides N., Purnelle B., Prozzi D.,
RA Ngui K., Masuy D., Hancy F., Burteau S., Boutry M., Delcours J.,
RA Goffeau A., Hols P.;
RT "Complete sequence and comparative genome analysis of the dairy
RT bacterium Streptococcus thermophilus.";
RL Nat. Biotechnol. 22:1554-1558(2004).
DR EMBL; CP000024; AAV61641.1; -; Genomic_DNA.
DR InterPro; IPR007921; CHAP.
DR InterPro; IPR009148; Siba.
DR Pfam; PF05257; CHAP; 1.
DR PRINTS; PR01852; SIBAPROTEIN.
DR PROSITE; PS50911; CHAP; 1.
KW Complete proteome.
SQ SEQUENCE 485 AA; 49165 MW; 8665E4B1FC6CF7CA CRC64;

Query Match 71.8%; Score 61; DB 2; Length 485;
Best Local Similarity 73.7%; Pred. No. 0.66;
Matches 14; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

QY 1 LSAVLVSGVTLSSAATLSA 19
||| ||||| ||| :|||
DB 6 LSAVLVSGVTLGAATTGVA 24

RESULT 14
Q9AKA4 STRAG
ID Q9AKA4_STRAG PRELIMINARY; PRT; 447 AA.
AC Q9AKA4;
DT 01-JUN-2001 (TrEMBLrel. 17, Created)
DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE PcsB protein precursor.
GN Name=pcsb;
OS Streptococcus agalactiae.
OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
OC Streptococcus.
OX NCBI_TaxID=1311;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=LMG 18311;
RX MEDLINE=21101799; PubMed=11157929;
RX DOI=10.1128/JB.183.4.1175-1183.2001;
RA Reinscheid D.J., Gottschalk B., Schubert A., Eikmanns B.J.,
RA Chhatwal G.S.;
RT "Identification and molecular analysis of PcsB, a protein required for
RT cell wall separation of group B streptococcus.";
RL J. Bacteriol. 183:1175-1183(2001).
DR EMBL; AJ277292; CAC28144.1; -; Genomic_DNA.
DR InterPro; IPR007921; CHAP.
DR InterPro; IPR009148; Siba.
DR Pfam; PF05257; CHAP; 1.
KW PRINTS; PR01852; SIBAPROTEIN.
SQ SEQUENCE 474 AA; 48142 MW; 1BAC6A9E0A0A200D CRC64;

Query Match 71.8%; Score 61; DB 2; Length 474;
Best Local Similarity 73.7%; Pred. No. 0.64;
Matches 14; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

QY 1 LSAVLVSGVTLSSAATLSA 19
||| ||||| ||| :|||

```

```

DB 6 LSAVLVSGVTLSSAASVHA 24

RESULT 13
Q5M212 STRT1
ID Q5M212_STRT1 PRELIMINARY; PRT; 485 AA.
AC Q5M212;
DT 01-FEB-2005 (TrEMBLrel. 29, Created)
DT 01-FEB-2005 (TrEMBLrel. 29, Last sequence update)
DT 01-FEB-2005 (TrEMBLrel. 29, Last annotation update)
DE Glucan binding protein.
GN Name=pcsb; OrderedLocusNames=str0022;
OS Streptococcus thermophilus (strain CNRZ 1056).
OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
OC Streptococcus.
OX NCBI_TaxID=299768;
RN [1]
RP NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
RX PubMed=15543133; DOI=10.1038/nbt1034;
RA Bolotin A., Quinkis B., Renault P., Sorokin A., Ehrlich S.D.,
RA Kulaukas S., Lapidus A., Goldsman E., Mazur M., Pusch G.D.,
RA Fonstein M., Overbeek R., Kyprides N., Purnelle B., Prozzi D.,
RA Ngui K., Masuy D., Hancy F., Burteau S., Boutry M., Delcours J.,
RA Goffeau A., Hols P.;
RT "Complete sequence and comparative genome analysis of the dairy
RT bacterium Streptococcus thermophilus.";
RL Nat. Biotechnol. 22:1554-1558(2004).
DR EMBL; CP000024; AAV61641.1; -; Genomic_DNA.
DR InterPro; IPR007921; CHAP.
DR InterPro; IPR009148; Siba.
DR Pfam; PF05257; CHAP; 1.
DR PRINTS; PR01852; SIBAPROTEIN.
DR PROSITE; PS50911; CHAP; 1.
KW Complete proteome.
SQ SEQUENCE 485 AA; 49165 MW; 8665E4B1FC6CF7CA CRC64;

Query Match 71.8%; Score 61; DB 2; Length 485;
Best Local Similarity 73.7%; Pred. No. 0.66;
Matches 14; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

QY 1 LSAVLVSGVTLSSAATLSA 19
||| ||||| ||| :|||
DB 6 LSAVLVSGVTLGAATTGVA 24

RESULT 14
Q9AKA4 STRAG
ID Q9AKA4_STRAG PRELIMINARY; PRT; 447 AA.
AC Q9AKA4;
DT 01-JUN-2001 (TrEMBLrel. 17, Created)
DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE PcsB protein precursor.
GN Name=pcsb;
OS Streptococcus agalactiae.
OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
OC Streptococcus.
OX NCBI_TaxID=1311;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=6313;
RX MEDLINE=21101799; PubMed=11157929;
RX DOI=10.1128/JB.183.4.1175-1183.2001;
RA Reinscheid D.J., Gottschalk B., Schubert A., Eikmanns B.J.,
RA Chhatwal G.S.;
RT "Identification and molecular analysis of PcsB, a protein required for
RT cell wall separation of group B streptococcus.";
RL J. Bacteriol. 183:1175-1183(2001).
DR EMBL; AJ277292; CAC28144.1; -; Genomic_DNA.
DR InterPro; IPR007921; CHAP.
DR InterPro; IPR009148; Siba.
DR Pfam; PF05257; CHAP; 1.
KW PRINTS; PR01852; SIBAPROTEIN.
SQ SEQUENCE 474 AA; 48142 MW; 1BAC6A9E0A0A200D CRC64;

Query Match 71.8%; Score 61; DB 2; Length 474;
Best Local Similarity 73.7%; Pred. No. 0.64;
Matches 14; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

QY 1 LSAVLVSGVTLSSAATLSA 19
||| ||||| ||| :|||

```

```

DR PROSITE; PS50911; CHAP; 1.
KW Signal.
FT SIGNAL 1 25 Potential.
FT CHAIN 26 447 PcsB protein.
SQ SEQUENCE 447 AA; 46681 MW; F4DB14B0A5F962C8 CRC64;

Query Match 62.9%; Score 53.5; DB 2; Length 447;
Best Local Similarity 70.0%; Pred. No. 9.1;
Matches 14; Conservative 4; Mismatches 1; Indels 1; Gaps 1;

QY 1 LSAVLVSGVTLSA 19
Db 6 LSAVLVSGVTLSA 25
|||||
|||||

RESULT 15
Q8E2H1 STRA5
ID Q8E2H1_STRAS PRELIMINARY; PRT; 447 AA.
AC Q8E2H1
DT 01-MAR-2003 (TrEMBLrel. 23, Created)
DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE PcsB protein.
GN Name=pcsB; OrderedLocusNames=SAG0017;
OS Streptococcus agalactiae (serotype V).
OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
OC Streptococcus.
OX NCBI_TaxID=216466;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=2603 V/R / Serotype V;
RX MEDLINE=22222988; PubMed=12200547; DOI=10.1073/pnas.182380799;
RA Tettelin H., Masiognani V., Cieleswicz M.J., Eisen J.A., Peterson S.N.,
RA Wesels M.R., Paulsen I.T., Nelson K.E., Margalit I., Read T.D.,
RA Madoff L.C., Wolf A.M., Beanan M.J., Brinkac L.M., Daugherty S.C.,
RA DeBoy R.T., Durkin A.S., Kolonay J.F., Madupu R., Lewis M.R.,
RA Radune D., Fedorova N.B., Scanlan D., Khouri H.M., Mulligan S.,
RA Carly H.A., Cline R.T., Van Aken S.E., Gill J., Scarselli M., Mora M.,
RA Iacobini E.T., Brettoni C., Galli G., Mariani M., Vegni F., Maione D.,
RA Rinaldo D., Rappuoli R., Telford J.L., Kasper D.L., Grandi G.,
RA Fraser C.M.;
RT "Complete genome sequence and comparative genomic analysis of an
RT emerging human pathogen, serotype V Streptococcus agalactiae.";
RL Proc. Natl. Acad. Sci. U.S.A. 99:12391-12396(2002).
DR EMBL; AE014192; NAM98925.1; -; Genomic_DNA.
DR TIGR; SAG0017; -.
DR InterPro; IPR007921; CHAP.
DR InterPro; IPR009148; Siba.
DR Pfam; PF05257; CHAP; 1.
DR PRINTS; PR01852; SIBAPROTEIN.
DR PROSITE; PS50911; CHAP; 1.
KW Complete proteome.
SQ SEQUENCE 447 AA; 46681 MW; F4DB14B0A5F962C8 CRC64;

Query Match 62.9%; Score 53.5; DB 2; Length 447;
Best Local Similarity 70.0%; Pred. No. 9.1;
Matches 14; Conservative 4; Mismatches 1; Indels 1; Gaps 1;

QY 1 LSAVLVSGVTLSA 19
Db 6 LSAVLVSGVTLSA 25
|||||
|||||

Search completed: January 25, 2006, 19:13:37
Job time : 72.3 secs

```

GenCore version 5.1.6
Copyright (c) 1993 - 2006 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: January 25, 2006, 18:50:05 ; Search time 16.6 Seconds
(without alignments)
99.609 Million cell updates/sec

Title: US-10-797-821-4
Perfect score: 85
Sequence: 1 LSAVLVSGVTLLSSATTL SAV 20

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 572060 seqs, 82675679 residues

Total number of hits satisfying chosen parameters: 572060

Minimum DB seq length: 0

Maximum DB seq length: 200000000
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 s

```
Database :
1: /cgm2_6/pdata/1/iaa/5_COMB.pep.*
2: /cgm2_6/pdata/1/iaa/6_COMB.pep.*
3: /cgm2_6/pdata/1/iaa/H_COMB.pep.*
4: /cgm2_6/pdata/1/iaa/PTCRS_COMB.pep.*
5: /cgm2_6/pdata/1/iaa/RE_COMB.pep.*
6: /cgm2_6/pdata/1/iaa/backfiles1.pep.*
```

SUMMARIES

Result No.	Query			DB	ID	Description
	Score	Match	%			
1	47	55.3	27	2	US-09-060-878-3	Sequence 3, Appli
2	47	55.3	92	1	US-08-186-222-4	Sequence 4, Appli
3	47	55.3	461	1	US-08-186-222-2	Sequence 2, Appli
4	46	54.1	363	2	US-09-543-681A-7064	Sequence 7064, Ap
5	44	51.8	73	2	US-09-621-976-7132	Sequence 7132, Ap
6	43	50.6	118	2	US-09-583-110-4129	Sequence 4129, Ap
7	43	50.6	120	2	US-09-732-210-655	Sequence 655, App
8	43	50.6	127	2	US-09-107-532A-5060	Sequence 5060, Ap
9	43	50.6	320	2	US-09-107-433-3472	Sequence 3472, Ap
10	43	50.6	320	1	US-08-362-739-2	Sequence 2, Appli
11	43	50.6	320	2	US-08-914-350A-2	Sequence 2, Appli
12	43	50.6	660	2	US-09-583-110-3976	Sequence 3976, Ap
13	42	49.4	182	2	US-10-622-064-2	Sequence 2, Appli
14	42	49.4	485	2	US-09-502-540-13623	Sequence 12623, A
15	42	49.4	486	2	US-09-198-452A-114	Sequence 114, App
16	42	49.4	486	2	US-09-438-185A-99	Sequence 99, Appl
17	42	49.4	539	2	US-09-252-991A-31813	Sequence 31813, A
18	42	49.4	555	2	US-09-270-767-45727	Sequence 45727, A
19	41	48.2	335	2	US-09-134-000C-3742	Sequence 3742, Ap
20	40	47.1	641	1	US-08-873-479-42	Sequence 42, Appl
21	40	47.1	824	2	US-08-873-479-42	Sequence 32329, A
22	40	47.1	1091	6	US-16232-991A-32329	Patent No. 5516630
23	40	47.1	2227	2	US-08-475-886-2	Sequence 2, Appli
24	40	47.1	2227	2	US-08-475-886-4	Sequence 4, Appli
25	40	47.1	2227	2	US-08-475-886-6	Sequence 6, Appli
26	40	47.1	2227	2	US-08-397-232-2	Sequence 2, Appli
27	40	47.1	2227	2	US-08-397-232-4	Sequence 4, Appli

28	40	47.1	2227	2	US-09-171-387-2	Sequence 2, Appli
29	40	47.1	2227	2	US-09-653-489-2	Sequence 2, Appli
30	40	47.1	2227	2	US-09-653-499-4	Sequence 4, Appli
31	40	47.1	2227	2	US-09-653-498-6	Sequence 6, Appli
32	40	47.1	2227	2	US-10-135-988-2	Sequence 2, Appli
33	40	47.1	2227	2	US-10-135-988-4	Sequence 4, Appli
34	40	47.1	2227	2	US-10-135-988-6	Sequence 6, Appli
35	39.5	46.5	703	2	US-09-134-000C-5072	Sequence 61372, Ap
36	39	45.9	71	2	US-09-270-767-61335	Sequence 61335, A
37	39	45.9	125	2	US-09-252-991A-3028	Sequence 3028, A
38	39	45.9	154	2	US-09-252-991A-18030	Sequence 18030, A
39	39	45.9	182	2	US-09-492-709A-361	Sequence 361, App
40	39	45.9	213	2	US-09-270-767-43368	Sequence 43368, A
41	39	45.9	256	2	US-09-270-767-45809	Sequence 45809, A
42	39	45.9	395	2	US-09-252-991A-18511	Sequence 18511, A
43	39	45.9	497	2	US-09-134-000C-5990	Sequence 5990, Ap
44	39	45.9	1305	2	US-09-715-962-6	Sequence 6, Appli
45	39	45.9	1356	2	US-09-770-170-6	Sequence 6, Appli

ALIGNMENTS

```

RESULT 1
US-09-060-878-3
; Sequence 3, Application US/09060878A
; Patent No. 6605286
; GENERAL INFORMATION:
; APPLICANT: Steidler, Lothar
; APPLICANT: Remaut, Erik
; APPLICANT: Wells, Jeremy Mark
; APPLICANT: Le Page, Richard William Falla
; TITLE OF INVENTION: Delivery of Biologically Active
; TITLE OF INVENTION: Polypeptides
; FILE REFERENCE: 8865-009-999
; CURRENT APPLICATION NUMBER: US/09/060,878A
; CURRENT FILING DATE: 1998-04-16
; EARLIER APPLICATION NUMBER: GB9521568.7
; EARLIER FILING DATE: 1995-10-20
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 3
; LENGTH: 27
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Leader Sequence of Usp45 in Lactococcus lactis
; Patent No. 6605286
; US-09-060-878-3

```

```
;
; COUNTRY: USA
; ZIP: 10532
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/186,222
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/672,205
; FILING DATE: 19-MAR-1991
; APPLICATION NUMBER: GB 9006400.7
; FILING DATE: 22-MAR-1990
; ATTORNEY/AGENT INFORMATION:
; NAME: Villamizar, Joann
; REGISTRATION NUMBER: 30,598
; REFERENCE/DOCKET NUMBER: 4-17994/A
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (914)785-7121
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 92 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-186-222-4

Query Match 55.3%; Score 47; DB 1; Length 92;
Best Local Similarity 55.0%; Pred. No. 2.2;
Matches 11; Conservative 4; Mismatches 5; Indels 0; Gaps 0;

QY 1 LSAVLVSGVTLSATLSAV 20
   :|:|:|:|:|:|:|:|:|
Db 6 ISAILMSTVILSAAAPLSGV 25

RESULT 3
US-08-186-222-2
; Sequence 2, Application US/08186222
; Patent No. 5559007
; GENERAL INFORMATION:
; APPLICANT: Suri, Bruno
; APPLICANT: Schmitz, Albert
; TITLE OF INVENTION: Bacterial Vectors
; NUMBER OF SEQUENCES: 4
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: CIBA-GEIGY Corporation
; STREET: 7 Skyline Drive
; CITY: Hawthorne
; STATE: New York
; COUNTRY: USA
; ZIP: 10532
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/186,222
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/672,205
; FILING DATE: 19-MAR-1991
; APPLICATION NUMBER: GB 9006400.7
; FILING DATE: 22-MAR-1990
; ATTORNEY/AGENT INFORMATION:
; NAME: Villamizar, Joann
; REGISTRATION NUMBER: 30,598

;
; COUNTRY: USA
; ZIP: 10532
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/186,222
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/672,205
; FILING DATE: 19-MAR-1991
; APPLICATION NUMBER: GB 9006400.7
; FILING DATE: 22-MAR-1990
; ATTORNEY/AGENT INFORMATION:
; NAME: Villamizar, Joann
; REGISTRATION NUMBER: 30,598

;
; REFERENCE/DOCKET NUMBER: 4-17994/A
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (914)785-7121
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 461 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-186-222-2

Query Match 55.3%; Score 47; DB 1; Length 461;
Best Local Similarity 55.0%; Pred. No. 16;
Matches 11; Conservative 4; Mismatches 5; Indels 0; Gaps 0;

QY 1 LSAVLVSGVTLSATLSAV 20
   :|:|:|:|:|:|:|:|:|
Db 6 ISAILMSTVILSAAAPLSGV 25

RESULT 4
US-09-543-681A-7064
; Sequence 7064, Application US/09543681A
; Patent No. 6605709
; GENERAL INFORMATION:
; APPLICANT: GARY BRETON
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PROTEUS MIRABILIS
; FILE REFERENCE: 2709.1002-001
; CURRENT APPLICATION NUMBER: US/09/543,681A
; CURRENT FILING DATE: 2000-04-05
; PRIOR APPLICATION NUMBER: US 60/128,706
; PRIOR FILING DATE: 1999-04-09
; NUMBER OF SEQ ID NOS: 8344
; SEQ ID NO 7064
; LENGTH: 363
; TYPE: PRT
; ORGANISM: Proteus mirabilis
; US-09-543-681A-7064

Query Match 54.1%; Score 46; DB 2; Length 363;
Best Local Similarity 52.6%; Pred. No. 17;
Matches 10; Conservative 3; Mismatches 6; Indels 0; Gaps 0;

QY 1 LSAVLVSGVTLSATLSA 19
   :|:|:|:|:|:|:|:|:|
Db 12 LSTLIISGTFLLSITTAQA 30

RESULT 5
US-09-621-976-7132
; Sequence 7132, Application US/09621976
; Patent No. 6639063
; GENERAL INFORMATION:
; APPLICANT: Dumas Milne Edwards, J.B.
; APPLICANT: Jobert, S.
; APPLICANT: Giordano, J.Y.
; TITLE OF INVENTION: ESTs and Encoded Human Proteins.
; FILE REFERENCE: GENSET.054PR2
; CURRENT APPLICATION NUMBER: US/09/621,976
; CURRENT FILING DATE: 2000-07-21
; NUMBER OF SEQ ID NOS: 19335
; SOFTWARE: Patent.pm
; SEQ ID NO 7132
; LENGTH: 73
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: UNSURE
; LOCATION: 70
; OTHER INFORMATION: Xaa = * ,Cys,Phe,Leu,Ser,Trp,Tyr
; NAME/KEY: UNSURE
```


; LOCATION: 73
 ; OTHER INFORMATION: Xaa = * ,Lys,Gln,Arg
 ; NAME/KEY: UNSURE
 ; LOCATION: 6
 ; OTHER INFORMATION: Xaa = Cys,Gly
 ; NAME/KEY: UNSURE
 ; LOCATION: 72
 ; OTHER INFORMATION: Xaa = Cys,Tyr
 ; NAME/KEY: UNSURE
 ; LOCATION: 69
 ; OTHER INFORMATION: Xaa = Ile,Lys,Arg,Thr
 ; NAME/KEY: UNSURE
 ; LOCATION: 71
 ; OTHER INFORMATION: Xaa = Leu,Met,Val
 US-09-621-976-7132

Query Match 51.8%; Score 44; DB 2; Length 73;
 Best Local Similarity 69.2%; Pred. No. 5;
 Matches 9; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 1 LSAVLVSGVTLS 13
 | :|||||:|
 Db 49 LESVLVSGITVSS 61

RESULT 6
 US-09-583-110-4129
 ; Sequence 4129, Application US/09583110
 ; Patent No. 6699703
 ; GENERAL INFORMATION:
 ; APPLICANT: Lynn Doucette-Stamm et al.
 ; TITLE OF INVENTION: Nucleic Acid and Amino Acid Sequences Relating to Streptococcus
 ; FILE REFERENCE: PATH00-07A
 ; CURRENT APPLICATION NUMBER: US/09/583,110
 ; CURRENT FILING DATE: 2000-05-26
 ; PRIOR APPLICATION NUMBER: US 09/107,433
 ; PRIOR FILING DATE: 1998-06-30
 ; PRIOR APPLICATION NUMBER: US 60/085,131
 ; PRIOR FILING DATE: 1998-05-12
 ; PRIOR APPLICATION NUMBER: US 60/051,553
 ; PRIOR FILING DATE: 1997-07-02
 ; NUMBER OF SEQ ID NOS: 5322
 ; SEQ ID NO 4129
 ; TYPE: PRT
 ; ORGANISM: Streptococcus pneumoniae
 US-09-583-110-4129

Query Match 50.6%; Score 43; DB 2; Length 118;
 Best Local Similarity 75.0%; Pred. No. 13;
 Matches 9; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

QY 6 VSGVTLSATTL 17
 | :|||||:|
 Db 48 VAGVTLASATL 59

RESULT 7
 US-09-732-210-655
 ; Sequence 655, Application US/09732210
 ; Patent No. 6573361
 ; GENERAL INFORMATION:
 ; APPLICANT: Bunkers, Greg J.
 ; APPLICANT: Liang, Jihong
 ; APPLICANT: Mittanck, Cindy A.
 ; APPLICANT: Seale, Jeffrey W.
 ; APPLICANT: Wu, Yomie S.
 ; TITLE OF INVENTION: Anti-fungal Proteins and Methods for Their Use
 ; FILE REFERENCE: 38-21(15036)B
 ; CURRENT APPLICATION NUMBER: US/09/732,210
 ; CURRENT FILING DATE: 2000-12-07
 ; PRIOR APPLICATION NUMBER: US 60/169,513

; PRIOR FILING DATE: 1999-12-07
 ; PRIOR APPLICATION NUMBER: US 60/169,340
 ; PRIOR FILING DATE: 1999-12-07
 ; NUMBER OF SEQ ID NOS: 1753
 ; SEQ ID NO 655
 ; LENGTH: 120
 ; TYPE: PRT
 ; ORGANISM: Bacillus subtilis
 US-09-732-210-655

Query Match 50.6%; Score 43; DB 2; Length 120;
 Best Local Similarity 75.0%; Pred. No. 13;
 Matches 9; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

QY 6 VSGVTLSATTL 17
 | :|||||:|
 Db 48 VNGVTLASATL 59

RESULT 8
 US-09-107-532A-5060
 ; Sequence 5060, Application US/09107532A
 ; Patent No. 6583275
 ; GENERAL INFORMATION:
 ; APPLICANT: Lynn A Doucette-Stamm and David Bush
 ; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO
 ; ENTEROCOCCUS FAECIUM FOR DIAGNOSTICS AND THERAPEUTICS
 ; NUMBER OF SEQUENCES: 7310
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSER: GENOME THERAPEUTICS CORPORATION
 ; STREET: 100 Beaver Street
 ; CITY: Waltham
 ; STATE: Massachusetts
 ; COUNTRY: USA
 ; ZIP: 02354
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: CD/ROM ISO9660
 ; COMPUTER: PC
 ; OPERATING SYSTEM: <Unknown>
 ; SOFTWARE: ASCII
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/09/107,532A
 ; FILING DATE: 30-Jun-1998
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER: 60/085,598
 ; FILING DATE: 14 May 1998
 ; APPLICATION NUMBER: 60/051571
 ; FILING DATE: July 2, 1997
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: Ariniello, Pamela Deneke
 ; REGISTRATION NUMBER: 40,489
 ; REFERENCE/DOCKET NUMBER: GTC-012
 ; TELECOMMUNICATION INFORMATION:
 ; TELEPHONE: (781)893-5007
 ; TELEFAX: (781)893-8277
 ; INFORMATION FOR SEQ ID NO: 5060:
 ; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 127 amino acids
 ; TYPE: amino acid
 ; TOPOLOGY: linear
 ; MOLECULE TYPE: protein
 ; HYPOTHETICAL: YES
 ; ORIGINAL SOURCE:
 ; ORGANISM: Enterococcus faecium
 ; FEATURE:
 ; NAME/KEY: misc.feature
 ; LOCATION: (B) LOCATION 1...127
 ; SEQUENCE DESCRIPTION: SEQ ID NO: 5060:

US-09-107-532A-5060
 ; Sequence 5060, Application US/09107532A
 ; Patent No. 6583275
 ; GENERAL INFORMATION:
 ; APPLICANT: Lynn A Doucette-Stamm and David Bush
 ; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO
 ; ENTEROCOCCUS FAECIUM FOR DIAGNOSTICS AND THERAPEUTICS
 ; NUMBER OF SEQUENCES: 7310
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSER: GENOME THERAPEUTICS CORPORATION
 ; STREET: 100 Beaver Street
 ; CITY: Waltham
 ; STATE: Massachusetts
 ; COUNTRY: USA
 ; ZIP: 02354
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: CD/ROM ISO9660
 ; COMPUTER: PC
 ; OPERATING SYSTEM: <Unknown>
 ; SOFTWARE: ASCII
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/09/107,532A
 ; FILING DATE: 30-Jun-1998
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER: 60/085,598
 ; FILING DATE: 14 May 1998
 ; APPLICATION NUMBER: 60/051571
 ; FILING DATE: July 2, 1997
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: Ariniello, Pamela Deneke
 ; REGISTRATION NUMBER: 40,489
 ; REFERENCE/DOCKET NUMBER: GTC-012
 ; TELECOMMUNICATION INFORMATION:
 ; TELEPHONE: (781)893-5007
 ; TELEFAX: (781)893-8277
 ; INFORMATION FOR SEQ ID NO: 5060:
 ; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 127 amino acids
 ; TYPE: amino acid
 ; TOPOLOGY: linear
 ; MOLECULE TYPE: protein
 ; HYPOTHETICAL: YES
 ; ORIGINAL SOURCE:
 ; ORGANISM: Enterococcus faecium
 ; FEATURE:
 ; NAME/KEY: misc.feature
 ; LOCATION: (B) LOCATION 1...127
 ; SEQUENCE DESCRIPTION: SEQ ID NO: 5060:

US-09-107-532A-5060
 ; Sequence 5060, Application US/09107532A
 ; Patent No. 6583275
 ; GENERAL INFORMATION:
 ; APPLICANT: Lynn A Doucette-Stamm and David Bush
 ; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO
 ; ENTEROCOCCUS FAECIUM FOR DIAGNOSTICS AND THERAPEUTICS
 ; NUMBER OF SEQUENCES: 7310
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSER: GENOME THERAPEUTICS CORPORATION
 ; STREET: 100 Beaver Street
 ; CITY: Waltham
 ; STATE: Massachusetts
 ; COUNTRY: USA
 ; ZIP: 02354
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: CD/ROM ISO9660
 ; COMPUTER: PC
 ; OPERATING SYSTEM: <Unknown>
 ; SOFTWARE: ASCII
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/09/107,532A
 ; FILING DATE: 30-Jun-1998
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER: 60/085,598
 ; FILING DATE: 14 May 1998
 ; APPLICATION NUMBER: 60/051571
 ; FILING DATE: July 2, 1997
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: Ariniello, Pamela Deneke
 ; REGISTRATION NUMBER: 40,489
 ; REFERENCE/DOCKET NUMBER: GTC-012
 ; TELECOMMUNICATION INFORMATION:
 ; TELEPHONE: (781)893-5007
 ; TELEFAX: (781)893-8277
 ; INFORMATION FOR SEQ ID NO: 5060:
 ; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 127 amino acids
 ; TYPE: amino acid
 ; TOPOLOGY: linear
 ; MOLECULE TYPE: protein
 ; HYPOTHETICAL: YES
 ; ORIGINAL SOURCE:
 ; ORGANISM: Enterococcus faecium
 ; FEATURE:
 ; NAME/KEY: misc.feature
 ; LOCATION: (B) LOCATION 1...127
 ; SEQUENCE DESCRIPTION: SEQ ID NO: 5060:

US-09-107-532A-5060
 ; Sequence 5060, Application US/09107532A
 ; Patent No. 6583275
 ; GENERAL INFORMATION:
 ; APPLICANT: Lynn A Doucette-Stamm and David Bush
 ; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO
 ; ENTEROCOCCUS FAECIUM FOR DIAGNOSTICS AND THERAPEUTICS
 ; NUMBER OF SEQUENCES: 7310
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSER: GENOME THERAPEUTICS CORPORATION
 ; STREET: 100 Beaver Street
 ; CITY: Waltham
 ; STATE: Massachusetts
 ; COUNTRY: USA
 ; ZIP: 02354
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: CD/ROM ISO9660
 ; COMPUTER: PC
 ; OPERATING SYSTEM: <Unknown>
 ; SOFTWARE: ASCII
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/09/107,532A
 ; FILING DATE: 30-Jun-1998
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER: 60/085,598
 ; FILING DATE: 14 May 1998
 ; APPLICATION NUMBER: 60/051571
 ; FILING DATE: July 2, 1997
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: Ariniello, Pamela Deneke
 ; REGISTRATION NUMBER: 40,489
 ; REFERENCE/DOCKET NUMBER: GTC-012
 ; TELECOMMUNICATION INFORMATION:
 ; TELEPHONE: (781)893-5007
 ; TELEFAX: (781)893-8277
 ; INFORMATION FOR SEQ ID NO: 5060:
 ; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 127 amino acids
 ; TYPE: amino acid
 ; TOPOLOGY: linear
 ; MOLECULE TYPE: protein
 ; HYPOTHETICAL: YES
 ; ORIGINAL SOURCE:
 ; ORGANISM: Enterococcus faecium
 ; FEATURE:
 ; NAME/KEY: misc.feature
 ; LOCATION: (B) LOCATION 1...127
 ; SEQUENCE DESCRIPTION: SEQ ID NO: 5060:

US-09-107-532A-5060
 ; Sequence 5060, Application US/09107532A
 ; Patent No. 6583275
 ; GENERAL INFORMATION:
 ; APPLICANT: Lynn A Doucette-Stamm and David Bush
 ; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO
 ; ENTEROCOCCUS FAECIUM FOR DIAGNOSTICS AND THERAPEUTICS
 ; NUMBER OF SEQUENCES: 7310
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSER: GENOME THERAPEUTICS CORPORATION
 ; STREET: 100 Beaver Street
 ; CITY: Waltham
 ; STATE: Massachusetts
 ; COUNTRY: USA
 ; ZIP: 02354
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: CD/ROM ISO9660
 ; COMPUTER: PC
 ; OPERATING SYSTEM: <Unknown>
 ; SOFTWARE: ASCII
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/09/107,532A
 ; FILING DATE: 30-Jun-1998
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER: 60/085,598
 ; FILING DATE: 14 May 1998
 ; APPLICATION NUMBER: 60/051571
 ; FILING DATE: July 2, 1997
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: Ariniello, Pamela Deneke
 ; REGISTRATION NUMBER: 40,489
 ; REFERENCE/DOCKET NUMBER: GTC-012
 ; TELECOMMUNICATION INFORMATION:
 ; TELEPHONE: (781)893-5007
 ; TELEFAX: (781)893-8277
 ; INFORMATION FOR SEQ ID NO: 5060:
 ; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 127 amino acids
 ; TYPE: amino acid
 ; TOPOLOGY: linear
 ; MOLECULE TYPE: protein
 ; HYPOTHETICAL: YES
 ; ORIGINAL SOURCE:
 ; ORGANISM: Enterococcus faecium
 ; FEATURE:
 ; NAME/KEY: misc.feature
 ; LOCATION: (B) LOCATION 1...127
 ; SEQUENCE DESCRIPTION: SEQ ID NO: 5060:

Query Match 50.6%; Score 43; DB 2; Length 127;
 Best Local Similarity 75.0%; Pred. No. 14;
 Matches 9; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

QY 6 VSGVTLSATL 17
|:||||:|:
Db 57 VAGVTLASATL 68

RESULT 9

US-09-107-433-3472
; Sequence 3472, Application US/09107433
; Patent No. 6800744

GENERAL INFORMATION:

APPLICANT: Lynn A Doucette-Stamm and David Bush
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID
SEQUENCES RELATING TO STREPTOCOCCUS PNEUMONIAE
THERAPEUTICS

NUMBER OF SEQUENCES: 5206

CORRESPONDENCE ADDRESS:

ADDRESSEE: GENOME THERAPEUTICS CORPORATION

STREET: 100 Beaver Street

CITY: Waltham

STATE: Massachusetts

COUNTRY: USA

ZIP: 02354

COMPUTER READABLE FORM:

MEDIUM TYPE: CD-ROM ISO9660

COMPUTER: <Unknown>

OPERATING SYSTEM: <Unknown>

SOFTWARE: <Unknown>

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/09/107,433

FILING DATE: 30-Jun-1998

PRIOR APPLICATION DATA:

APPLICATION NUMBER: 60/ 085131

FILING DATE: May 12, 1998

APPLICATION NUMBER: 60/051553

FILING DATE: July 2, 1997

ATTORNEY/AGENT INFORMATION:

NAME: Ariniello, Pamela Deneke

REGISTRATION NUMBER: 40,489

REFERENCE/DOCKET NUMBER: GTC-011

TELECOMMUNICATION INFORMATION:

TELEPHONE: (781)893-5007

TELEFAX: (781)893-8277

INFORMATION FOR SEQ ID NO: 3472:

SEQUENCE CHARACTERISTICS:

LENGTH: 140 amino acids

TYPE: amino acid

TOPOLOGY: linear

MOLECULE TYPE: protein

HYPOTHETICAL: YES

ORIGINAL SOURCE:

ORGANISM: Streptococcus pneumoniae

FEATURE:

NAME/KEY: misc_feature

LOCATION: (B) LOCATION 1...140

SEQUENCE DESCRIPTION: SEQ ID NO: 3472:

US-09-107-433-3472

Query Match 50.6%; Score 43; DB 2; Length 140;
Best Local Similarity 75.0%; Pred. No. 16;
Matches 9; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

QY 6 VSGVTLSATL 17
|:||||:|:
Db 70 VAGVTLASATL 81

RESULT 10

US-08-362-739-2
; Sequence 2, Application US/08362739

; Patent No. 5658757

; GENERAL INFORMATION:

; APPLICANT: Haake, David A.

; APPLICANT: Blanco, David R.
; APPLICANT: Champion, Cheryl I.
; APPLICANT: Lovett, Michael A.
; APPLICANT: Miller, James N.
; TITLE OF INVENTION: CLONED Leptosira OUTER MEMBRANE PROTEIN
; NUMBER OF SEQUENCES: 7
; CORRESPONDENCE ADDRESS:

ADDRESSEE: Spensley Horn Jubas & Lubitz
STREET: 1880 Century Park East, Suite 500
CITY: Los Angeles
STATE: California
COUNTRY: USA
ZIP: 90067

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: PatentIn Release #1.0, Version #1.25

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/362,739

FILING DATE:

CLASSIFICATION: 435

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 08/040,747

FILING DATE: 31-MAR-1993

ATTORNEY/AGENT INFORMATION:

NAME: Wetherell, Jr. Ph.D., John R.,

REGISTRATION NUMBER: 31,678

REFERENCE/DOCKET NUMBER: PD-2097

TELECOMMUNICATION INFORMATION:

TELEPHONE: (619) 455-5100

TELEFAX: (619) 455-5110

INFORMATION FOR SEQ ID NO: 2:

SEQUENCE CHARACTERISTICS:

LENGTH: 320 amino acids

TYPE: amino acid

TOPOLOGY: linear

MOLECULE TYPE: protein

US-08-362-739-2

Query Match 50.6%; Score 43; DB 1; Length 320;
Best Local Similarity 58.8%; Pred. No. 45;
Matches 10; Conservative 3; Mismatches 4; Indels 0; Gaps 0;

QY 3 AVLVSGVTLSATLSA 19
|:|:|:|:|:|:
Db 8 ALLILAVLSSAASLSA 24

RESULT 11

US-08-914-350A-2

; Sequence 2, Application US/08914350A

; Patent No. 6685945

; GENERAL INFORMATION:

; APPLICANT: THE REGENTS OF THE UNIVERSITY OF CALIFORNIA

; APPLICANT: HAAKE, David A.

; APPLICANT: BLANCO, David R.

; APPLICANT: CHAMPION, Cheryl I.

; APPLICANT: LOVETT, Michael A.

; APPLICANT: MILLER, James N.

; TITLE OF INVENTION: CLONED LEPTOSPIRA OUTER MEMBRANE PROTEIN

; FILE REFERENCE: UCLA1420-2

; CURRENT APPLICATION NUMBER: US/08/914,350A

; CURRENT FILING DATE: 1997-08-19

; PRIOR APPLICATION NUMBER: US 08/362,739

; PRIOR FILING DATE: 1994-12-20

; PRIOR APPLICATION NUMBER: US 08/040,747

; PRIOR FILING DATE: 1993-03-31

; NUMBER OF SEQ ID NOS: 7

; SOFTWARE: PatentIn version 3.1

; SEQ ID NO 2

; LENGTH: 320

; TYPE: PRT

6 TATVVI.SAISI.SSTTAA 2A

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2006 Compugen Ltd.

OM protein - protein search, using sw model

Run on: January 25, 2006, 19:13:51 ; Search time 60 Seconds
(without alignments)
139.276 Million cell updates/sec

Title: US-10-797-821-4
Perfect score: 85
Sequence: 1 LSAVLVSGVTLSSATLSAV 20

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1867569 seqs, 417829326 residues

Total number of hits satisfying chosen parameters: 1867569

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published Applications AA Main:
1: /cgn2_6/prodata/1/pubpaa/US07_PUBCOMB.pep.*
2: /cgn2_6/prodata/1/pubpaa/US08_PUBCOMB.pep.*
3: /cgn2_6/prodata/1/pubpaa/US09_PUBCOMB.pep.*
4: /cgn2_6/prodata/1/pubpaa/US10A_PUBCOMB.pep.*
5: /cgn2_6/prodata/1/pubpaa/US10B_PUBCOMB.pep.*
6: /cgn2_6/prodata/1/pubpaa/US11_PUBCOMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Length	DB ID	Description
1	85	100.0	20	4	US-10-383-930-4
2	85	100.0	20	5	US-10-797-821-4
3	85	100.0	431	4	US-10-383-930-29
4	85	100.0	431	4	US-10-383-930-30
5	85	100.0	431	4	US-10-383-930-33
6	85	100.0	431	5	US-10-797-821-29
7	85	100.0	431	5	US-10-797-821-30
8	85	100.0	431	5	US-10-797-821-33
9	85	100.0	432	4	US-10-383-930-32
10	85	100.0	432	5	US-10-797-821-32
11	84	98.8	432	4	US-10-383-930-31
12	84	98.8	432	5	US-10-797-821-31
13	67	78.8	398	5	US-10-474-792-600
14	48	56.5	177	4	US-10-425-114-47654
15	47	55.3	27	3	US-09-060-878-3
16	47	55.3	27	4	US-10-350-250-3
17	47	55.3	32	4	US-10-431-577-63
18	46	54.1	657	4	US-10-425-115-251444
19	43	50.6	69	5	US-10-472-928-565
20	43	50.6	118	4	US-10-282-122A-57579
21	43	50.6	118	4	US-10-282-122A-72213
22	43	50.6	118	4	US-10-282-122A-73739
23	43	50.6	118	4	US-10-282-122A-74331
24	43	50.6	118	4	US-10-474-776-608
25	43	50.6	118	5	US-10-472-928-268
26	43	50.6	122	5	US-10-501-282-390
27	43	50.6	131	3	US-09-815-242-13214

28	43	50.6	131	3	US-09-815-242-13572	Sequence 13572, A
29	43	50.6	140	5	US-10-617-320-3472	Sequence 3472, Ap
30	43	50.6	172	4	US-10-425-115-364037	Sequence 364037, A
31	43	50.6	195	4	US-10-282-122A-74812	Sequence 74812, A
32	43	50.6	201	4	US-10-425-115-285996	Sequence 285996, A
33	43	50.6	320	2	US-08-914-350-2	Sequence 2, Appli
34	43	50.6	320	4	US-10-322-391-2	Sequence 2, Appli
35	43	50.6	798	5	US-10-450-763-53091	Sequence 53091, A
36	42	49.4	20	4	US-10-383-930-5	Sequence 5, Appli
37	42	49.4	20	5	US-10-797-821-5	Sequence 5, Appli
38	42	49.4	182	3	US-09-916-230-8	Sequence 8, Appli
39	42	49.4	182	3	US-09-848-616-146	Sequence 146, App
40	42	49.4	182	4	US-10-289-454-8	Sequence 8, Appli
41	42	49.4	182	4	US-10-050-902-146	Sequence 146, App
42	42	49.4	182	4	US-10-050-898-146	Sequence 146, App
43	42	49.4	182	4	US-10-346-190-8	Sequence 8, Appli
44	42	49.4	182	4	US-10-289-456-8	Sequence 8, Appli
45	42	49.4	182	4	US-10-622-064-2	Sequence 2, Appli

ALIGNMENTS

RESULT 1
US-10-383-930-4
; Sequence 4, Application US/10383930
; Publication NO. US20040127400A1
; GENERAL INFORMATION:
; APPLICANT: Smith, Daniel J
; APPLICANT: Taubman, Martin A
; TITLE OF INVENTION: Immunogenicity of Glucan Binding Protein
; FILE REFERENCE: 25669-018
; CURRENT APPLICATION NUMBER: US/10/383,930
; CURRENT FILING DATE: 2003-03-07
; PRIOR APPLICATION NUMBER: 60/402,483
; PRIOR FILING DATE: 2002-08-08
; PRIOR APPLICATION NUMBER: 60/363,209
; PRIOR FILING DATE: 2002-03-07
; NUMBER OF SEQ ID NOS: 41
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 4
; LENGTH: 20
; TYPE: PRT
; ORGANISM: Streptococcus mutans
US-10-383-930-4
Query Match 100.0%; Score 85; DB 4; Length 20;
Best Local Similarity 100.0%; Pred.No. 2.4e-06;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 LSAVLVSGVTLSSATLSAV 20
|||||
Db 1 LSAVLVSGVTLSSATLSAV 20
RESULT 2
US-10-797-821-4
; Sequence 4, Application US/10797821
; Publication No. US20050031633A1
; GENERAL INFORMATION:
; APPLICANT: Smith, Daniel J.
; APPLICANT: Taubman, Martin A.
; TITLE OF INVENTION: Glucan Binding Protein and Glycosyltransferase Immunogens
; FILE REFERENCE: 25669-020
; CURRENT APPLICATION NUMBER: US/10/797,821
; CURRENT FILING DATE: 2004-03-09
; PRIOR APPLICATION NUMBER: 10/383,930
; PRIOR FILING DATE: 2003-03-07
; PRIOR APPLICATION NUMBER: 60/363,209
; PRIOR FILING DATE: 2002-03-07
; PRIOR APPLICATION NUMBER: 60/402,483
; PRIOR FILING DATE: 2002-08-08
; PRIOR APPLICATION NUMBER: 09/290,049

; PRIOR FILING DATE: 1999-04-12
; PRIOR APPLICATION NUMBER: 60/081,550
; PRIOR FILING DATE: 1998-04-13
; PRIOR APPLICATION NUMBER: 60/115,142
; PRIOR FILING DATE: 1999-01-08
; NUMBER OF SEQ ID NOS: 45
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 4
; LENGTH: 20
; TYPE: PRT
; ORGANISM: Artificial
; FEATURE:
; OTHER INFORMATION: GbpB peptide
US-10-797-821-4

Query Match 100.0%; Score 85; DB 5; Length 20;
Best Local Similarity 100.0%; Pred. No. 2.4e-06; Indels 0; Gaps 0;
Matches 20; Conservative 0; Mismatches 0;

Qy 1 LSAVLVSGVTLSSTLSSAV 20
|||||
Db 1 LSAVLVSGVTLSSTLSSAV 20

RESULT 3
US-10-383-930-29
; Sequence 29, Application US/10383930
; Publication No. US20040127400A1
; GENERAL INFORMATION:
; APPLICANT: Smith, Daniel J
; TITLE OF INVENTION: Immunogenicity of Glucan Binding Protein
; FILE REFERENCE: 25669-018
; CURRENT APPLICATION NUMBER: US/10/383,930
; CURRENT FILING DATE: 2003-03-07
; PRIOR APPLICATION NUMBER: 60/402,483
; PRIOR FILING DATE: 2002-08-08
; PRIOR APPLICATION NUMBER: 60/363,209
; PRIOR FILING DATE: 2002-03-07
; NUMBER OF SEQ ID NOS: 41
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 29
; LENGTH: 431
; TYPE: PRT
; ORGANISM: Streptococcus mutans
US-10-383-930-29

Query Match 100.0%; Score 85; DB 4; Length 431;
Best Local Similarity 100.0%; Pred. No. 8.8e-05; Indels 0; Gaps 0;
Matches 20; Conservative 0; Mismatches 0;

Qy 1 LSAVLVSGVTLSSTLSSAV 20
|||||
Db 6 LSAVLVSGVTLSSTLSSAV 25

RESULT 4
US-10-383-930-30
; Sequence 30, Application US/10383930
; Publication No. US20040127400A1
; GENERAL INFORMATION:
; APPLICANT: Smith, Daniel J
; TITLE OF INVENTION: Immunogenicity of Glucan Binding Protein
; FILE REFERENCE: 25669-018
; CURRENT APPLICATION NUMBER: US/10/383,930
; CURRENT FILING DATE: 2003-03-07
; PRIOR APPLICATION NUMBER: 60/402,483
; PRIOR FILING DATE: 2002-08-08
; PRIOR APPLICATION NUMBER: 60/363,209
; PRIOR FILING DATE: 2002-03-07
; NUMBER OF SEQ ID NOS: 41
; SOFTWARE: PatentIn version 3.2

; SEQ ID NO 30
; LENGTH: 431
; TYPE: PRT
; ORGANISM: Streptococcus mutans
US-10-383-930-30

Query Match 100.0%; Score 85; DB 4; Length 431;
Best Local Similarity 100.0%; Pred. No. 8.8e-05; Indels 0; Gaps 0;
Matches 20; Conservative 0; Mismatches 0;

Qy 1 LSAVLVSGVTLSSTLSSAV 20
|||||
Db 6 LSAVLVSGVTLSSTLSSAV 25

RESULT 5
US-10-383-930-33
; Sequence 33, Application US/10383930
; Publication No. US20040127400A1
; GENERAL INFORMATION:
; APPLICANT: Smith, Daniel J
; TITLE OF INVENTION: Immunogenicity of Glucan Binding Protein
; FILE REFERENCE: 25669-018
; CURRENT APPLICATION NUMBER: US/10/383,930
; CURRENT FILING DATE: 2003-03-07
; PRIOR APPLICATION NUMBER: 60/402,483
; PRIOR FILING DATE: 2002-08-08
; PRIOR APPLICATION NUMBER: 60/363,209
; PRIOR FILING DATE: 2002-03-07
; NUMBER OF SEQ ID NOS: 41
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 33
; LENGTH: 431
; TYPE: PRT
; ORGANISM: Streptococcus mutans
US-10-383-930-33

Query Match 100.0%; Score 85; DB 4; Length 431;
Best Local Similarity 100.0%; Pred. No. 8.8e-05; Indels 0; Gaps 0;
Matches 20; Conservative 0; Mismatches 0;

Qy 1 LSAVLVSGVTLSSTLSSAV 20
|||||
Db 6 LSAVLVSGVTLSSTLSSAV 25

RESULT 6
US-10-797-821-29
; Sequence 29, Application US/10797821
; Publication No. US20050031633A1
; GENERAL INFORMATION:
; APPLICANT: Taubman, Martin A.
; TITLE OF INVENTION: Glucan Binding Protein and Glycosyltransferase Immunogens
; FILE REFERENCE: 25669-020
; CURRENT APPLICATION NUMBER: US/10/797,821
; CURRENT FILING DATE: 2004-03-09
; PRIOR APPLICATION NUMBER: 10/383,930
; PRIOR FILING DATE: 2003-03-07
; PRIOR APPLICATION NUMBER: 60/363,209
; PRIOR FILING DATE: 2002-03-07
; PRIOR APPLICATION NUMBER: 60/402,483
; PRIOR FILING DATE: 2002-08-08
; PRIOR APPLICATION NUMBER: 09/290,049
; PRIOR FILING DATE: 1999-04-12
; PRIOR APPLICATION NUMBER: 60/081,550
; PRIOR FILING DATE: 1998-04-13
; PRIOR APPLICATION NUMBER: 60/115,142
; PRIOR FILING DATE: 1999-01-08
; NUMBER OF SEQ ID NOS: 45
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 29

; LENGTH: 431
 ; TYPE: PRT
 ; ORGANISM: Streptococcus mutans
 US-10-797-821-29

Query Match 100.0%; Score 85; DB 5; Length 431;
 Best Local Similarity 100.0%; Pred. No. 8.8e-05;
 Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 LSAVLVSGVTLSSATTL SAV 20
 Db 6 LSAVLVSGVTLSSATTL SAV 25

RESULT 7
 US-10-797-821-30
 ; Sequence 30, Application US/10797821
 ; Publication No. US20050031633A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Smith, Daniel J.
 ; APPLICANT: Taubman, Martin A.
 ; TITLE OF INVENTION: Glucan Binding Protein and Glycosyltransferase Immunogens
 ; FILE REFERENCE: 25669-020
 ; CURRENT APPLICATION NUMBER: US/10/797,821
 ; CURRENT FILING DATE: 2004-03-09
 ; PRIOR APPLICATION NUMBER: 10/383,930
 ; PRIOR FILING DATE: 2003-03-07
 ; PRIOR APPLICATION NUMBER: 60/363,209
 ; PRIOR FILING DATE: 2002-03-07
 ; PRIOR APPLICATION NUMBER: 60/402,483
 ; PRIOR FILING DATE: 2002-08-08
 ; PRIOR APPLICATION NUMBER: 09/290,049
 ; PRIOR FILING DATE: 1999-04-12
 ; PRIOR APPLICATION NUMBER: 60/081,550
 ; PRIOR FILING DATE: 1998-04-13
 ; PRIOR APPLICATION NUMBER: 60/115,142
 ; PRIOR FILING DATE: 1999-01-08
 ; NUMBER OF SEQ ID NOS: 45
 ; SOFTWARE: Patent in version 3.2
 ; SEQ ID NO 30
 ; LENGTH: 431
 ; TYPE: PRT
 ; ORGANISM: Streptococcus mutans
 US-10-797-821-30

Query Match 100.0%; Score 85; DB 5; Length 431;
 Best Local Similarity 100.0%; Pred. No. 8.8e-05;
 Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 LSAVLVSGVTLSSATTL SAV 20
 Db 6 LSAVLVSGVTLSSATTL SAV 25

RESULT 8
 US-10-797-821-33
 ; Sequence 33, Application US/10797821
 ; Publication No. US20050031633A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Smith, Daniel J.
 ; APPLICANT: Taubman, Martin A.
 ; TITLE OF INVENTION: Glucan Binding Protein and Glycosyltransferase Immunogens
 ; FILE REFERENCE: 25669-020
 ; CURRENT APPLICATION NUMBER: US/10/797,821
 ; CURRENT FILING DATE: 2004-03-09
 ; PRIOR APPLICATION NUMBER: 10/383,930
 ; PRIOR FILING DATE: 2003-03-07
 ; PRIOR APPLICATION NUMBER: 60/363,209
 ; PRIOR FILING DATE: 2002-03-07
 ; PRIOR APPLICATION NUMBER: 60/402,483
 ; PRIOR FILING DATE: 2002-08-08
 ; PRIOR APPLICATION NUMBER: 09/290,049
 ; PRIOR FILING DATE: 1999-04-12

; PRIOR APPLICATION NUMBER: 60/081,550
 ; PRIOR FILING DATE: 1998-04-13
 ; PRIOR APPLICATION NUMBER: 60/115,142
 ; PRIOR FILING DATE: 1999-01-08
 ; NUMBER OF SEQ ID NOS: 45
 ; SOFTWARE: Patent in version 3.2
 ; SEQ ID NO 33
 ; LENGTH: 431
 ; TYPE: PRT
 ; ORGANISM: Streptococcus mutans
 US-10-797-821-33

Query Match 100.0%; Score 85; DB 5; Length 431;
 Best Local Similarity 100.0%; Pred. No. 8.8e-05;
 Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 LSAVLVSGVTLSSATTL SAV 20
 Db 6 LSAVLVSGVTLSSATTL SAV 25

RESULT 9
 US-10-383-930-32
 ; Sequence 32, Application US/10383930
 ; Publication No. US20040127400A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Smith, Daniel J.
 ; APPLICANT: Taubman, Martin A.
 ; TITLE OF INVENTION: Immunogenicity of Glucan Binding Protein
 ; FILE REFERENCE: 25669-018
 ; CURRENT APPLICATION NUMBER: US/10/383,930
 ; CURRENT FILING DATE: 2003-03-07
 ; PRIOR APPLICATION NUMBER: 60/402,483
 ; PRIOR FILING DATE: 2002-08-08
 ; PRIOR APPLICATION NUMBER: 60/363,209
 ; PRIOR FILING DATE: 2002-03-07
 ; NUMBER OF SEQ ID NOS: 41
 ; SOFTWARE: Patent in version 3.2
 ; SEQ ID NO 32
 ; LENGTH: 432
 ; TYPE: PRT
 ; ORGANISM: Streptococcus mutans
 US-10-383-930-32

Query Match 100.0%; Score 85; DB 4; Length 432;
 Best Local Similarity 100.0%; Pred. No. 8.9e-05;
 Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 LSAVLVSGVTLSSATTL SAV 20
 Db 6 LSAVLVSGVTLSSATTL SAV 25

RESULT 10
 US-10-797-821-32
 ; Sequence 32, Application US/10797821
 ; Publication No. US20050031633A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Smith, Daniel J.
 ; APPLICANT: Taubman, Martin A.
 ; TITLE OF INVENTION: Glucan Binding Protein and Glycosyltransferase Immunogens
 ; FILE REFERENCE: 25669-020
 ; CURRENT APPLICATION NUMBER: US/10/797,821
 ; CURRENT FILING DATE: 2004-03-09
 ; PRIOR APPLICATION NUMBER: 10/383,930
 ; PRIOR FILING DATE: 2003-03-07
 ; PRIOR APPLICATION NUMBER: 60/363,209
 ; PRIOR FILING DATE: 2002-03-07
 ; PRIOR APPLICATION NUMBER: 60/402,483
 ; PRIOR FILING DATE: 2002-08-08
 ; PRIOR APPLICATION NUMBER: 09/290,049
 ; PRIOR FILING DATE: 1999-04-12

Fri Jan 27 10:12:19 2006

```
; PRIOR FILING DATE: 1998-04-13
; PRIOR APPLICATION NUMBER: 60/115,142
; PRIOR FILING DATE: 1999-01-08
; NUMBER OF SEQ ID NOS: 45
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 31
; LENGTH: 432
; TYPE: PRT
; ORGANISM: Streptococcus mutans
US-10-797-821-32

Query Match      100.0%; Score 85; DB 5; Length 432;
Best Local Similarity 100.0%; Pred. No. 8.9e-05;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 LSAVLVSGVTLSSATLSAV 20
   |||||
Db 6 LSAVLVSGVTLSSATLSAI 25

RESULT 11
US-10-383-930-31
; Sequence 31, Application US/10383930
; Publication No. US20040127400A1
; GENERAL INFORMATION:
; APPLICANT: Smith, Daniel J
; APPLICANT: Taubman, Martin A
; TITLE OF INVENTION: Immunogenicity of Glucan Binding Protein
; FILE REFERENCE: 25669-018
; CURRENT APPLICATION NUMBER: US/10/383,930
; CURRENT FILING DATE: 2003-03-07
; PRIOR APPLICATION NUMBER: 60/402,483
; PRIOR FILING DATE: 2002-08-08
; PRIOR APPLICATION NUMBER: 60/363,209
; PRIOR FILING DATE: 2002-03-07
; NUMBER OF SEQ ID NOS: 41
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 31
; LENGTH: 432
; TYPE: PRT
; ORGANISM: Streptococcus mutans
US-10-383-930-31

Query Match      98.8%; Score 84; DB 4; Length 432;
Best Local Similarity 95.0%; Pred. No. 0.00013;
Matches 19; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 LSAVLVSGVTLSSATLSAV 20
   |||||
Db 6 LSAVLVSGVTLSSATLSAI 25

RESULT 12
US-10-797-821-31
; Sequence 31, Application US/10797821
; Publication No. US20050031633A1
; GENERAL INFORMATION:
; APPLICANT: Smith, Daniel J.
; APPLICANT: Taubman, Martin A.
; TITLE OF INVENTION: Glucan Binding Protein and Glycosyltransferase Immunogens
; FILE REFERENCE: 25669-020
; CURRENT APPLICATION NUMBER: US/10/797,821
; CURRENT FILING DATE: 2004-03-09
; PRIOR APPLICATION NUMBER: 10/383,930
; PRIOR FILING DATE: 2003-03-07
; PRIOR APPLICATION NUMBER: 60/363,209
; PRIOR FILING DATE: 2002-03-07
; PRIOR APPLICATION NUMBER: 60/402,483
; PRIOR FILING DATE: 2002-08-08
; PRIOR APPLICATION NUMBER: 09/290,049
; PRIOR FILING DATE: 1999-04-12
; PRIOR APPLICATION NUMBER: 60/081,550
; PRIOR FILING DATE: 1998-04-13

; PRIOR FILING DATE: 1998-04-13
; PRIOR APPLICATION NUMBER: 60/115,142
; PRIOR FILING DATE: 1999-01-08
; NUMBER OF SEQ ID NOS: 45
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 31
; LENGTH: 432
; TYPE: PRT
; ORGANISM: Streptococcus mutans
US-10-797-821-31

Query Match      98.8%; Score 84; DB 5; Length 432;
Best Local Similarity 95.0%; Pred. No. 0.00013;
Matches 19; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 LSAVLVSGVTLSSATLSAV 20
   |||||
Db 6 LSAVLVSGVTLSSATLSAI 25

RESULT 13
US-10-474-792-600
; Sequence 600, Application US/10474792
; Publication No. US20040236072A1
; GENERAL INFORMATION:
; APPLICANT: Olmsted, Stephen
; APPLICANT: Zagursky, Robert
; APPLICANT: Nickbarg, Elliot
; APPLICANT: Winter, Lourie
; TITLE OF INVENTION: SURFACE PROTEINS OF STREPTOCOCCUS PYOGENES
; FILE REFERENCE: AM 100399
; CURRENT APPLICATION NUMBER: US/10/474,792
; CURRENT FILING DATE: 2003-10-14
; NUMBER OF SEQ ID NOS: 674
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 600
; LENGTH: 398
; TYPE: PRT
; ORGANISM: Streptococcus pyogenes
US-10-474-792-600

Query Match      78.8%; Score 67; DB 5; Length 398;
Best Local Similarity 78.9%; Pred. No. 0.052;
Matches 15; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

Qy 1 LSAVLVSGVTLSSATLSA 19
   |||||
Db 6 LSAVLVSGVTLGAATTGGA 24

RESULT 14
US-10-425-114-47654
; Sequence 47654, Application US/10425114
; Publication No. US20040034888A1
; GENERAL INFORMATION:
; APPLICANT: Liu, Jingdong
; APPLICANT: Zhou, Yihua
; APPLICANT: Kovalic, David K.
; APPLICANT: Screen, Steven E
; APPLICANT: Tabaska, Jack E
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated with
; FILE REFERENCE: 38-21(53313)B
; CURRENT APPLICATION NUMBER: US/10/425,114
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 73128
; SEQ ID NO 47654
; LENGTH: 177
; TYPE: PRT
; ORGANISM: Glycine max
; FEATURE:
; OTHER INFORMATION: Clone ID: LIB3051-080-B6_FLI.pap
US-10-425-114-47654
```


Query Match 56.5%; Score 48; DB 4; Length 177;
 Best Local Similarity 60.0%; Pred. No. 18;
 Matches 12; Conservative 1; Mismatches 7; Indels 0; Gaps 0;

QY 1 LSAVLVSGVTLSSATTLSAV 20
 |||||:|||||
 Db 80 LSAVLVAANTLKERTASAV 99

RESULT 15
 US-09-060-878-3
 ; Sequence 3, Application US/09060878A
 ; Patent No. US20010006642A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Steidl, Erik
 ; APPLICANT: Remaut, Erik
 ; APPLICANT: Wells, Jeremy Mark
 ; APPLICANT: Le Page, Richard William Falla
 ; TITLE OF INVENTION: Delivery of Biologically Active
 ; FILE REFERENCE: 8865-009-999
 ; CURRENT APPLICATION NUMBER: US/09/060,878A
 ; EARLIER FILING DATE: 1998-04-16
 ; EARLIER APPLICATION NUMBER: GB9521568.7
 ; EARLIER FILING DATE: 1995-10-20
 ; NUMBER OF SEQ ID NOS: 6
 ; SOFTWARE: FastSeq for Windows Version 3.0
 ; SEQ ID NO 3
 ; LENGTH: 27
 ; TYPE: PRT
 ; ORGANISM: Artificial Sequence
 ; FEATURE:
 ; OTHER INFORMATION: Leader Sequence of Usp45 in Lactococcus lactis
 US-09-060-878-3

Query Match 55.3%; Score 47; DB 3; Length 27;
 Best Local Similarity 55.0%; Pred. No. 2.9;
 Matches 11; Conservative 4; Mismatches 5; Indels 0; Gaps 0;

QY 1 LSAVLVSGVTLSSATTLSAV 20
 :|||:|:|:|:|:|
 Db 6 ISAILMSTVILSAAPLSGV 25

Search completed: January 25, 2006, 20:10:02
 Job time : 61 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2006 Compugen Ltd.

OM protein - protein search, using sw model

Run on: January 25, 2006, 19:15:26 ; Search time 5.5 Seconds
(without alignments)
39.378 Million cell updates/sec

Title: US-10-797-821-4
Perfect score: 85
Sequence: 1 LSAVLVSGVTLSSATLSAV 20

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 75621 seqs, 10829074 residues

Total number of hits satisfying chosen parameters: 75621

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published Applications AA New:
1: /cgn2_6/prodata/2/pubpaa/US08_NEW_PUB.pep.*
2: /cgn2_6/prodata/2/pubpaa/US06_NEW_PUB.pep.*
3: /cgn2_6/prodata/2/pubpaa/US07_NEW_PUB.pep.*
4: /cgn2_6/prodata/2/pubpaa/PCT_NEW_PUB.pep.*
5: /cgn2_6/prodata/2/pubpaa/US09_NEW_PUB.pep.*
6: /cgn2_6/prodata/2/pubpaa/US10_NEW_PUB.pep.*
7: /cgn2_6/prodata/2/pubpaa/US11_NEW_PUB.pep.*
8: /cgn2_6/prodata/2/pubpaa/US60_NEW_PUB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	85	100.0	431	7	US-11-052-554A-210 Sequence 210, App
2	67	78.8	398	7	US-11-052-554A-252 Sequence 252, App
3	47	55.3	29	7	US-11-021-441-46 Sequence 46, Appl
4	45	52.9	313	7	US-11-156-084-352 Sequence 352, App
5	42	49.4	182	7	US-11-125-402-2 Sequence 2, Appli
6	39	45.9	106	6	US-10-485-788A-833 Sequence 833, App
7	39	45.9	106	7	US-11-053-076-218 Sequence 218, App
8	39	45.9	124	7	US-11-052-554A-333 Sequence 333, App
9	38	44.7	200	6	US-10-793-626-2126 Sequence 2126, Ap
10	38	44.7	414	6	US-10-793-626-806 Sequence 806, App
11	38	44.7	443	6	US-10-467-657-5762 Sequence 5762, Ap
12	38	44.7	823	6	US-10-467-657-2526 Sequence 2526, Ap
13	38	44.7	865	7	US-11-094-519A-41 Sequence 41, Appl
14	37	43.5	182	7	US-11-052-554A-25 Sequence 25, Appl
15	37	43.5	182	7	US-11-052-554A-181 Sequence 181, App
16	37	43.5	199	7	US-11-053-185-28 Sequence 28, Appl
17	37	43.5	334	6	US-10-467-657-1034 Sequence 1034, Ap
18	37	43.5	37	6	US-10-763-712A-75 Sequence 75, Appl
19	37	43.5	1007	6	US-10-517-939-84 Sequence 84, Appl
20	36.5	42.9	1237	7	US-11-052-554A-95 Sequence 95, Appl
21	36	42.4	135	7	US-11-108-172-203 Sequence 203, App
22	36	42.4	880	7	US-11-171-701-2 Sequence 2, Appli
23	35	41.2	23	7	US-11-058-384-6 Sequence 6, Appli
24	35	41.2	61	7	US-11-147-994-3 Sequence 3, Appli
25	35	41.2	61	7	US-11-147-994-5 Sequence 5, Appli

26	35	41.2	141	7	US-11-103-957-51	Sequence 51, Appl
27	35	41.2	297	6	US-10-453-372-70	Sequence 70, Appl
28	35	41.2	412	7	US-11-082-389-96	Sequence 96, Appl
29	35	41.2	428	6	US-10-513-639-17	Sequence 17, Appl
30	35	41.2	550	6	US-10-453-372-76	Sequence 76, Appl
31	35	41.2	578	6	US-10-453-372-66	Sequence 66, Appl
32	35	41.2	578	6	US-10-453-372-80	Sequence 80, Appl
33	35	41.2	600	6	US-10-878-556A-43	Sequence 385, App
34	35	41.2	660	7	US-11-033-039-385	Sequence 13, Appl
35	35	41.2	661	7	US-11-155-288-13	Sequence 1, Appli
36	35	41.2	661	7	US-11-119-502-1	Sequence 9, Appli
37	35	41.2	662	7	US-11-090-439-9	Sequence 22, Appl
38	35	41.2	780	7	US-11-089-551A-22	Sequence 3, Appli
39	35	41.2	1016	7	US-11-103-957-41	Sequence 4, Appli
40	35	41.2	1162	6	US-10-451-375-3	Sequence 283, App
41	35	41.2	1198	6	US-10-451-375-4	Sequence 1262, Ap
42	35	41.2	1461	7	US-11-052-554A-283	Sequence 141, App
43	35	41.2	2657	6	US-10-821-234-1262	Sequence 72, Appl
44	35	41.2	3716	7	US-11-052-554A-141	
45	34.5	40.6	348	7	US-11-129-143-72	

ALIGNMENTS

RESULT 1
US-11-052-554A-210
; Sequence 210, Application US/11052554A
; Publication No. US20050288866A1
; GENERAL INFORMATION:
; APPLICANT: Sachdeva, et al.
; TITLE OF INVENTION: COMPUTATIONAL METHOD FOR IDENTIFYING ADHESIN AND ADHESIN-LIKE
; FILE REFERENCE: 30853/40359A
; CURRENT APPLICATION NUMBER: US/11/052,554A
; CURRENT FILING DATE: 2005-02-07
; PRIOR APPLICATION NUMBER: US 60/589,227
; PRIOR FILING DATE: 2004-07-20
; PRIOR APPLICATION NUMBER: IN 173/DEL/2004
; PRIOR FILING DATE: 2004-02-06
; NUMBER OF SEQ ID NOS: 763
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 210
; LENGTH: 431
; TYPE: PRT
; ORGANISM: Streptococcus mutans UAL159
US-11-052-554A-210

Query Match 100.0%; Score 85; DB 7; Length 431;
Best Local Similarity 100.0%; Pred. No. 3.6e-06;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 LSAVLVSGVTLSSATLSAV 20
Db 6 LSAVLVSGVTLSSATLSAV 25
|||||
RESULT 2
US-11-052-554A-252
; Sequence 252, Application US/11052554A
; Publication No. US20050288866A1
; GENERAL INFORMATION:
; APPLICANT: Sachdeva, et al.
; TITLE OF INVENTION: COMPUTATIONAL METHOD FOR IDENTIFYING ADHESIN AND ADHESIN-LIKE
; FILE REFERENCE: 30853/40359A
; CURRENT APPLICATION NUMBER: US/11/052,554A
; CURRENT FILING DATE: 2005-02-07
; PRIOR APPLICATION NUMBER: US 60/589,227
; PRIOR FILING DATE: 2004-07-20
; PRIOR APPLICATION NUMBER: IN 173/DEL/2004
; PRIOR FILING DATE: 2004-02-06
; NUMBER OF SEQ ID NOS: 763

```

; SEQ ID NO 352
; LENGTH: 313
; TYPE: PRT
; ORGANISM: Photorhabdus luminescens subsp. laumondii TTO1
US-11-156-084-352

Query Match          52.9%; Score 45; DB 7; Length 313;
Best Local Similarity 55.6%; Pred. No. 3.8;
Matches 10; Conservative 2; Mismatches 6; Indels 0; Gaps 0;

QY      1 LSAVLVSGVTLSATTLS 18
Db      172 LEVFLISGKLTTELTTLS 189

RESULT 5
US-11-125-402-2
; Sequence 2, Application US/11125402
; Publication No. US20050281845A1
; GENERAL INFORMATION:
; APPLICANT: Bachmann, Martin F
; APPLICANT: Maurer, Patrick F
; TITLE OF INVENTION: Hapten-Carrier Conjugates and Uses Thereof
; FILE REFERENCE: 1700.0300001
; CURRENT APPLICATION NUMBER: US/11/125,402
; CURRENT FILING DATE: 2005-05-10
; PRIOR APPLICATION NUMBER: US/10/622,064
; PRIOR FILING DATE: 2003-07-18
; PRIOR APPLICATION NUMBER: US 60/396,575
; PRIOR FILING DATE: 2002-07-18
; NUMBER OF SEQ ID NOS: 33
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 2
; LENGTH: 182
; TYPE: PRT
; ORGANISM: Escherichia coli
US-11-125-402-2

Query Match          49.4%; Score 42; DB 7; Length 182;
Best Local Similarity 47.4%; Pred. No. 5.9;
Matches 9; Conservative 6; Mismatches 4; Indels 0; Gaps 0;

QY      1 LSAVLVSGVTLSATTLSA 19
Db      6 LAIVLSALSSTTALAA 24

RESULT 6
US-10-485-788A-833
; Sequence 833, Application US/10485788A
; Publication No. US20050282743A1
; GENERAL INFORMATION:
; APPLICANT: Lu, Peter S.
; APPLICANT: Rabinowitz, Joshua D.
; APPLICANT: Schweizer, Johannes
; APPLICANT: Carrick, Deanna Marie
; APPLICANT: Arbor Vita Corporation
; TITLE OF INVENTION: Molecular Interactions in Cells
; FILE REFERENCE: 20054-003320US
; CURRENT APPLICATION NUMBER: US/10/485,788A
; CURRENT FILING DATE: 2004-02-03
; PRIOR APPLICATION NUMBER: US 60/309,841
; PRIOR FILING DATE: 2001-08-03
; PRIOR APPLICATION NUMBER: US 60/360,061
; PRIOR FILING DATE: 2002-02-25
; PRIOR APPLICATION NUMBER: WO PCT/US02/24655
; PRIOR FILING DATE: 2002-08-02
; NUMBER OF SEQ ID NOS: 841
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 833
; LENGTH: 106
; TYPE: PRT
; ORGANISM: Homo sapiens

```

US-10-485-788A-833

Query Match 45.9%; Score 39; DB 6; Length 106;
Best Local Similarity 41.2%; Pred. No. 9.2;
Matches 7; Conservative 7; Mismatches 3; Indels 0; Gaps 0;

QY 4 VLVSGVTLSSATLSAV 20
|:|:|:|:|:|:|:|:|:|:
Db 60 VVWGVSMENATSAFAI 76

RESULT 7

US-11-053-076-218
; Sequence 218, Application US/11053076
; Publication No. US20050255460A1
; GENERAL INFORMATION:
; APPLICANT: Lu, Peter S.
; APPLICANT: Schweizer, Johannes
; APPLICANT: Somoza Diaz-Sarmiento, Chamorro
; APPLICANT: Belmares, Michael P.
; TITLE OF INVENTION: METHODS OF DIAGNOSING CERVICAL CANCER
; FILE REFERENCE: VITA-008CIP
; CURRENT APPLICATION NUMBER: US/11/053,076
; CURRENT FILING DATE: 2005-02-07
; PRIOR APPLICATION NUMBER: PCT/US03/28508
; PRIOR FILING DATE: 2003-09-09
; PRIOR APPLICATION NUMBER: 10/630,590
; PRIOR FILING DATE: 2003-07-29
; PRIOR APPLICATION NUMBER: 60/490,094
; PRIOR FILING DATE: 2003-07-25
; PRIOR APPLICATION NUMBER: 60/450,464
; PRIOR FILING DATE: 2003-02-27
; PRIOR APPLICATION NUMBER: 60/409,298
; PRIOR FILING DATE: 2002-09-09
; PRIOR APPLICATION NUMBER: 10/630,590
; PRIOR FILING DATE: 2003-07-29
; PRIOR APPLICATION NUMBER: PCT/US02/24655
; PRIOR FILING DATE: 2002-08-02
; PRIOR APPLICATION NUMBER: 60/309,841
; PRIOR FILING DATE: 2001-08-03
; PRIOR APPLICATION NUMBER: 60/360,061
; PRIOR FILING DATE: 2002-02-25
; PRIOR APPLICATION NUMBER: 10/080,273
; PRIOR FILING DATE: 2002-02-19
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 330
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 218
; LENGTH: 106
; TYPE: PRT
; ORGANISM: Homo sapiens

US-11-053-076-218

Query Match 45.9%; Score 39; DB 7; Length 106;
Best Local Similarity 41.2%; Pred. No. 9.2;
Matches 7; Conservative 7; Mismatches 3; Indels 0; Gaps 0;

QY 4 VLVSGVTLSSATLSAV 20
|:|:|:|:|:|:|:|:|:|:
Db 60 VVWGVSMENATSAFAI 76

RESULT 8

US-11-052-554A-333
; Sequence 333, Application US/11052554A
; Publication No. US20050288666A1
; GENERAL INFORMATION:
; APPLICANT: Sachdeva, et al.
; TITLE OF INVENTION: COMPUTATIONAL METHOD FOR IDENTIFYING ADHESIN AND ADHESIN-LIKE
; FILE REFERENCE: 30853/40359A
; CURRENT APPLICATION NUMBER: US/11/052,554A
; CURRENT FILING DATE: 2005-02-07

; PRIOR APPLICATION NUMBER: US 60/589,227
; PRIOR FILING DATE: 2004-07-20
; PRIOR APPLICATION NUMBER: IN 173/DEL/2004
; PRIOR FILING DATE: 2004-02-06
; NUMBER OF SEQ ID NOS: 763
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 333
; LENGTH: 124
; TYPE: PRT
; ORGANISM: Mycobacterium tuberculosis H37RV
US-11-052-554A-333

Query Match 45.9%; Score 39; DB 7; Length 124;
Best Local Similarity 50.0%; Pred. No. 11;
Matches 8; Conservative 5; Mismatches 3; Indels 0; Gaps 0;

QY 3 AVLVSQVTLSSATLS 18
|:|:|:|:|:|:|:|:|:|:
Db 12 AVAVAGVATATTTV 27

RESULT 9

US-10-793-626-2126
; Sequence 2126, Application US/10793626
; Publication No. US20050255478A1
; GENERAL INFORMATION:
; APPLICANT: KIMMERLY, WILLIAM JOHN
; TITLE OF INVENTION: STAPHYLOCOCCUS EPIDERMIDIS NUCLEIC ACIDS AND PROTEINS
; FILE REFERENCE: PU3480US
; CURRENT APPLICATION NUMBER: US/10/793,626
; CURRENT FILING DATE: 2004-03-04
; PRIOR APPLICATION NUMBER: 60/164,258
; PRIOR FILING DATE: 1999-11-09
; NUMBER OF SEQ ID NOS: 4472
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 2126
; LENGTH: 200
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: synthetic
; OTHER INFORMATION: amino acid sequence
US-10-793-626-2126

Query Match 44.7%; Score 38; DB 6; Length 200;
Best Local Similarity 52.9%; Pred. No. 27;
Matches 9; Conservative 4; Mismatches 4; Indels 0; Gaps 0;

QY 4 VLVSGVTLSSATLSAV 20
|:|:|:|:|:|:|:|:|:|:
Db 47 VVVTGVTLSSENPPASV 63

RESULT 10

US-10-793-626-806
; Sequence 806, Application US/10793626
; Publication No. US20050255478A1
; GENERAL INFORMATION:
; APPLICANT: KIMMERLY, WILLIAM JOHN
; TITLE OF INVENTION: STAPHYLOCOCCUS EPIDERMIDIS NUCLEIC ACIDS AND PROTEINS
; FILE REFERENCE: PU3480US
; CURRENT APPLICATION NUMBER: US/10/793,626
; CURRENT FILING DATE: 2004-03-04
; PRIOR APPLICATION NUMBER: 60/164,258
; PRIOR FILING DATE: 1999-11-09
; NUMBER OF SEQ ID NOS: 4472
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 806
; LENGTH: 414
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: synthetic
US-10-793-626-806


```
; GENERAL INFORMATION:
; APPLICANT: Sachdeva, et al.
; TITLE OF INVENTION: COMPUTATIONAL METHOD FOR IDENTIFYING ADHESIN AND ADHESIN-LIKE
; TITLE OF INVENTION: PROTEINS OF THERAPEUTIC POTENTIAL
; FILE REFERENCE: 30853/40359A
; CURRENT APPLICATION NUMBER: US/11/052,554A
; CURRENT FILING DATE: 2005-02-07
; PRIOR APPLICATION NUMBER: US 60/589,227
; PRIOR FILING DATE: 2004-07-20
; PRIOR APPLICATION NUMBER: IN 173/DEL/2004
; PRIOR FILING DATE: 2004-02-06
; NUMBER OF SEQ ID NOS: 763
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 181
; LENGTH: 182
; TYPE: PRT
; ORGANISM: Shigella flexneri 2a str. 2457T
US-11-052-554A-181
```

```
Query Match          43.5%; Score 37; DB 7; Length 182;
Best Local Similarity 44.4%; Pred. No. 35;
Matches 8; Conservative 6; Mismatches 4; Indels 0; Gaps 0;
```

```
OY      1  LSAVLVSGVTLSSATLS 18
      | : | : | : | : | : | :
Db      6  LAIVLSALSLSAALA 23
```

Search completed: January 25, 2006, 20:11:07
Job time : 5.5 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2006 Compugen Ltd.

OM protein - protein search, using sw model

Run on: January 25, 2006, 18:48:24 ; Search time 68.2 Seconds
(without alignments)
128.850 Million cell updates/sec

Title: US-10-797-821-5
Perfect score: 92
Sequence: 1 LSSATTL SAVKADDPDAQIA 20

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 2443163 seqs, 439378781 residues

Total number of hits satisfying chosen parameters: 2443163

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : A_Geneseq_21.*
1: Geneseq1980s:*
2: Geneseq1990s:*
3: Geneseq2000s:*
4: Geneseq2001s:*
5: Geneseq2002s:*
6: Geneseq2003as:*
7: Geneseq2003bs:*
8: Geneseq2004s:*
9: Geneseq2005s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	92	100.0	20	7	ADD93625 Streptoco
2	92	100.0	20	9	ADX37248 Streptoco
3	92	100.0	431	7	ADD93649 Streptoco
4	92	100.0	431	7	ADD93650 Streptoco
5	92	100.0	431	7	ADD93653 Streptoco
6	92	100.0	431	9	ADX37272 Streptoco
7	92	100.0	431	9	ADX37273 Streptoco
8	92	100.0	431	9	ADX37276 Streptoco
9	92	100.0	431	9	AEb91500 Microbial
10	92	100.0	432	7	ADD93652 Streptoco
11	92	100.0	432	9	ADX37275 Streptoco
12	91	98.9	432	7	ADD93651 Streptoco
13	91	98.9	432	9	ADX37274 Streptoco
14	46	50.0	694	4	ABB70601 Drosophil
15	46	50.0	4643	5	ABB71609 Drosophil
16	45	48.9	447	5	ABP29684 Streptoco
17	45	48.9	447	8	ADU69524 S agalact
18	45	48.9	447	8	ADV98392 Streptoco
19	45	48.9	447	8	ADV81808 Streptoco
20	45	48.9	447	8	ADV79645 Streptoco
21	44	47.8	58	4	AM67505 Human imm
22	44	47.8	446	6	ABU23386 Protein e
23	44	47.8	1260	6	ADA35459 Acinetoba
24	43	46.7	380	8	ADN20492 Bacterial

25	43	46.7	483	8	ADX68803	Adx68803 Plant ful
26	43	46.7	524	6	ABU29734	Abu29734 Protein e
27	43	46.7	525	7	ADC95468	Adc95468 E. faeciu
28	43	46.7	1042	6	ABR57376	AbR57376 M. gallis
29	43	46.7	1062	2	AAW22721	Aaw22721 Antigenic
30	43	46.7	1224	6	ABR57375	AbR57375 Mycoplasma
31	43	46.7	1432	6	ABU24715	Abu24715 Protein e
32	42	45.7	20	7	ADD93624	Add93624 Streptoco
33	42	45.7	20	9	ADX37247	Adx37247 Streptoco
34	42	45.7	427	6	ABU15262	Abu15262 Protein e
35	42	45.7	427	8	ADN18116	Adn18116 Bacterial
36	42	45.7	497	5	ABP69600	Abp69600 Human pol
37	42	45.7	675	9	ABW95237	Abw95237 M. xanthu
38	41	44.6	127	4	ABM60826	Abm60826 Drosophil
39	41	44.6	167	5	ABB47487	Abb47487 Listeria
40	41	44.6	174	3	AAG23502	Aag23502 Arabidops
41	41	44.6	174	3	AAG39522	Aag39522 Arabidops
42	41	44.6	206	3	AAG39521	Aag39521 Arabidops
43	41	44.6	207	3	AAG23501	Aag23501 Arabidops
44	41	44.6	310	4	ABB62585	Abb62585 Drosophil
45	41	44.6	356	6	ABM67126	Abm67126 Photorhab

ALIGNMENTS

RESULT 1
ADD93625
ID ADD93625 standard; peptide; 20 AA.
XX
AC ADD93625;
XX
DT 29-JAN-2004 (first entry)
XX
DE Streptococcus mutans glucan binding protein-B peptide fragment.
XX
KW Glucan binding protein-B; GbpB; vaccine; anticaries; epitope; immunogen.
XX
OS Streptococcus mutans.
XX
PN WO2003075845-A2.
XX
PD 18-SEP-2003.
XX
PF 07-MAR-2003; 2003WO-US0006962.
XX
PR 07-MAR-2002; 2002US-0363209P.
PR 08-AUG-2002; 2002US-0402483P.
XX
(FORS-) FORSYTH INST.
Smith DJ, Taubman MA;
WPI; 2003-845091/78.
Composition useful as vaccines for dental caries comprises a fragment of a glucan binding protein-B binding to a major histocompatibility complex class II protein.
Claim 11; Page 10; 49pp; English.

The present sequence is that of a peptide comprising amino acid residues 16-35 of the glucan binding protein-B (GbpB) of Streptococcus mutans strain SK32 ADP93649. The peptide binds to a major histocompatibility complex (MHC) class II protein. It was identified as a potential B cell epitope using a matrix-based algorithm for epitope prediction, which was used to search the primary amino acid sequence of GbpB for known MHC class II binding motifs. The peptide can be used in immunogenic compositions and subunit vaccines for dental caries. These comprise MHC class II protein-binding GbpB peptides covalently linked with peptide subunits (preferably from the catalytic domain) of a glucosyltransferase. The compositions are used in a claimed method of eliciting production of an antibody in a mammal. Diepitopic or multi-epitopic polypeptides can be

CC prepared synthetically or by recombinant DNA technology. Antibodies
 CC raised against MHC class II binding fragments of GbpB can be used in
 CC passive immunisation.

SQ Sequence 20 AA;
 Query Match 100.0%; Score 92; DB 7; Length 20;
 Best Local Similarity 100.0%; Pred. No. 2.4e-08;
 Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 LSSATTLSAVKADDFDAQIA 20
 |||||
 Db 1 LSSATTLSAVKADDFDAQIA 20
 |||||

RESULT 2
 ADX37248
 ID ADX37248 standard; peptide; 20 AA.
 XX
 AC ADX37248;
 XX
 DT 21-APR-2005 (first entry)
 XX
 DE Streptococcus mutant glucan binding protein B peptide #5.
 XX
 KW immunogenicity; immune stimulation; glucan binding protein-B;
 KW microparticle; major histocompatibility complex; tooth disease.
 XX
 OS Streptococcus mutans.
 XX
 US2005031633-A1.
 XX
 PD 10-FEB-2005.
 XX
 PF 09-MAR-2004; 2004US-00797821.
 XX
 PR 13-APR-1998; 98US-0081550P.
 PR 08-JAN-1999; 99US-0115142P.
 PR 12-APR-1999; 99US-00290049.
 PR 07-MAR-2002; 2002US-0363209P.
 PR 08-AUG-2002; 2002US-0402483P.
 PR 07-MAR-2003; 2003US-00383930.
 XX
 PA (SMIT/) SMITH D J.
 PA (TAUB/) TAUBMAN M A.
 XX
 PI Smith DJ, Taubman MA;
 XX
 DR WPI; 2005-151644/16.
 XX

PT New composition comprising a fragment of a glucan binding protein-B
 PT (GbpB) that binds to MHC class II protein, and a biocompatible
 PT microparticle, useful for producing an antibody (claimed) for immunizing
 PT mammals against dental caries.
 XX
 PS Claim 4; SEQ ID NO 5; 73pp; English.
 XX
 CC The invention relates to a composition comprising a fragment of a glucan
 CC binding protein-B (GbpB) and a biocompatible microparticle, where the
 CC fragment binds to a major histocompatibility complex (MHC) class II
 CC protein. The composition is useful for producing an antibody for
 CC immunizing mammals against dental caries. This sequence corresponds to a
 CC Streptococcus mutans GbpB-derived peptide of the invention.
 XX
 SQ Sequence 20 AA;

Query Match 100.0%; Score 92; DB 9; Length 20;
 Best Local Similarity 100.0%; Pred. No. 2.4e-08;
 Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 LSSATTLSAVKADDFDAQIA 20
 |||||
 Db 1 LSSATTLSAVKADDFDAQIA 20
 |||||

RESULT 3
 ADD93649
 ID ADD93649 standard; protein; 431 AA.
 XX
 AC ADD93649;
 XX
 DT 29-JAN-2004 (first entry)
 XX
 DE Streptococcus mutans glucan binding protein-B.
 KW Glucan binding protein-B; GbpB; vaccine; anticaries; epitope; immunogen.
 XX
 OS Streptococcus mutans.
 XX
 FH Key Location/Qualifiers
 FT Region 6..25
 FT /note= "HLA-binding peptide"
 FT Region 16..35
 FT /note= "HLA-binding peptide"
 FT Region 33..52
 FT /note= "HLA-binding peptide"
 FT Region 37..56
 FT /note= "HLA-binding peptide"
 FT Region 48..67
 FT /note= "HLA-binding peptide"
 FT Region 52..71
 FT /note= "HLA-binding peptide"
 FT Region 88..107
 FT /note= "HLA-binding peptide"
 FT Region 113..132
 FT /note= "HLA-binding peptide"
 FT Region 117..136
 FT /note= "HLA-binding peptide"
 FT Region 137..156
 FT /note= "HLA-binding peptide"
 FT Region 174..193
 FT /note= "HLA-binding peptide"
 FT Region 194..213
 FT /note= "HLA-binding peptide"
 FT Region 214..233
 FT /note= "HLA-binding peptide"
 FT Region 248..267
 FT /note= "HLA-binding peptide"
 FT Region 289..308
 FT /note= "HLA-binding peptide"
 FT Region 306..325
 FT /note= "HLA-binding peptide"
 FT Region 311..330
 FT /note= "HLA-binding peptide"
 FT Region 349..368
 FT /note= "HLA-binding peptide"
 FT Region 365..384
 FT /note= "HLA-binding peptide"
 FT Region 383..402
 FT /note= "HLA-binding peptide"
 FT Region 403..422
 FT /note= "HLA-binding peptide"
 XX
 WO2003075845-A2.
 XX
 PD 18-SEP-2003.
 XX
 PF 07-MAR-2003; 2003WO-US006962.
 XX
 PR 07-MAR-2002; 2002US-0363209P.
 PR 08-AUG-2002; 2002US-0402483P.
 XX
 PA (FORS-) FORSYTH INST.
 XX
 PI Smith DJ, Taubman MA;
 XX

Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 LSSATTLGAVKADDFDAQIA 20
 Db 16 LSSATTLGAVKADDFDAQIA 35

RESULT 6
 ADX37272
 ID ADX37272 standard; protein; 431 AA.
 XX
 AC ADX37272;
 XX
 DT 21-APR-2005 (first entry)
 XX
 DE Streptococcus mutant glucan binding protein B variant #1.
 XX
 KW immunogenicity; immune stimulation; glucan binding protein-B;
 KW microparticle; major histocompatibility complex; tooth disease.
 XX
 OS Streptococcus mutans.
 XX
 PN US2005031633-A1.
 XX
 PD 10-FEB-2005.
 XX
 PF 09-MAR-2004; 2004US-00797821.
 XX
 PR 13-APR-1998; 98US-0081550P.
 PR 08-JAN-1999; 99US-0115142P.
 PR 12-APR-1999; 99US-00290049.
 PR 07-MAR-2002; 2002US-0363209P.
 PR 08-AUG-2002; 2002US-0402483P.
 PR 07-MAR-2003; 2003US-00383930.
 XX
 PA (SMIT/) SMITH D J.
 PA (TAUB/) TAUBMAN M A.
 XX
 PI Smith DJ, Taubman MA;
 XX
 PF WPI; 2005-151644/16.
 XX
 PR New composition comprising a fragment of a glucan binding protein-B (GbpB) that binds to MHC class II protein, and a biocompatible microparticle, useful for producing an antibody (claimed) for immunizing mammals against dental caries.
 PT
 PS Claim 3; SEQ ID NO 29; 73pp; English.
 XX
 CC The invention relates to a composition comprising a fragment of a glucan binding protein-B (GbpB) and a biocompatible microparticle, where the fragment binds to a major histocompatibility complex (MHC) class II protein. The composition is useful for producing an antibody for immunizing mammals against dental caries. This sequence corresponds to a Streptococcus mutans GbpB protein of the invention.
 CC
 XX
 SQ Sequence 431 AA;
 Query Match 100.0%; Score 92; DB 9; Length 431;
 Best Local Similarity 100.0%; Pred. No. 1e-06;
 Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 LSSATTLGAVKADDFDAQIA 20
 Db 16 LSSATTLGAVKADDFDAQIA 35

RESULT 7
 ADX37273
 ID ADX37273 standard; protein; 431 AA.
 XX
 AC ADX37273;
 XX
 DT 21-APR-2005 (first entry)
 XX
 DE Streptococcus mutant glucan binding protein B variant #5.
 XX
 KW immunogenicity; immune stimulation; glucan binding protein-B;
 KW microparticle; major histocompatibility complex; tooth disease.
 XX
 OS Streptococcus mutans.
 XX
 PN US2005031633-A1.
 XX
 PD 10-FEB-2005.
 XX
 PF 09-MAR-2004; 2004US-00797821.
 XX
 PR 13-APR-1998; 98US-0081550P.
 PR 08-JAN-1999; 99US-0115142P.
 PR 12-APR-1999; 99US-00290049.
 PR 07-MAR-2002; 2002US-0363209P.
 PR 08-AUG-2002; 2002US-0402483P.
 PR 07-MAR-2003; 2003US-00383930.
 XX
 PA (SMIT/) SMITH D J.
 PA (TAUB/) TAUBMAN M A.
 XX
 PI Smith DJ, Taubman MA;
 XX
 PF WPI; 2005-151644/16.
 XX
 PR New composition comprising a fragment of a glucan binding protein-B (GbpB) that binds to MHC class II protein, and a biocompatible microparticle, useful for producing an antibody (claimed) for immunizing mammals against dental caries.
 PT
 PS Claim 3; SEQ ID NO 29; 73pp; English.
 XX
 CC The invention relates to a composition comprising a fragment of a glucan binding protein-B (GbpB) and a biocompatible microparticle, where the fragment binds to a major histocompatibility complex (MHC) class II protein. The composition is useful for producing an antibody for immunizing mammals against dental caries. This sequence corresponds to a Streptococcus mutans GbpB protein of the invention.
 CC
 XX
 SQ Sequence 431 AA;
 Query Match 100.0%; Score 92; DB 9; Length 431;
 Best Local Similarity 100.0%; Pred. No. 1e-06;
 Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

DT 21-APR-2005 (first entry)
 XX
 DE Streptococcus mutant glucan binding protein B variant #2.
 XX
 KW immunogenicity; immune stimulation; glucan binding protein-B;
 KW microparticle; major histocompatibility complex; tooth disease.
 XX
 OS Streptococcus mutans.
 XX
 PN US2005031633-A1.
 XX
 PD 10-FEB-2005.
 XX
 PF 09-MAR-2004; 2004US-00797821.
 XX
 PR 13-APR-1998; 98US-0081550P.
 PR 08-JAN-1999; 99US-0115142P.
 PR 12-APR-1999; 99US-00290049.
 PR 07-MAR-2002; 2002US-0363209P.
 PR 08-AUG-2002; 2002US-0402483P.
 PR 07-MAR-2003; 2003US-00383930.
 XX
 PA (SMIT/) SMITH D J.
 PA (TAUB/) TAUBMAN M A.
 XX
 PI Smith DJ, Taubman MA;
 XX
 PF WPI; 2005-151644/16.
 XX
 PR New composition comprising a fragment of a glucan binding protein-B (GbpB) that binds to MHC class II protein, and a biocompatible microparticle, useful for producing an antibody (claimed) for immunizing mammals against dental caries.
 PT
 PS Claim 3; SEQ ID NO 30; 73pp; English.
 XX
 CC The invention relates to a composition comprising a fragment of a glucan binding protein-B (GbpB) and a biocompatible microparticle, where the fragment binds to a major histocompatibility complex (MHC) class II protein. The composition is useful for producing an antibody for immunizing mammals against dental caries. This sequence corresponds to a Streptococcus mutans GbpB protein of the invention.
 CC
 XX
 SQ Sequence 431 AA;
 Query Match 100.0%; Score 92; DB 9; Length 431;
 Best Local Similarity 100.0%; Pred. No. 1e-06;
 Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 LSSATTLGAVKADDFDAQIA 20
 Db 16 LSSATTLGAVKADDFDAQIA 35

RESULT 8
 ADX37276
 ID ADX37276 standard; protein; 431 AA.
 XX
 AC ADX37276;
 XX
 DT 21-APR-2005 (first entry)
 XX
 DE Streptococcus mutant glucan binding protein B variant #5.
 XX
 KW immunogenicity; immune stimulation; glucan binding protein-B;
 KW microparticle; major histocompatibility complex; tooth disease.
 XX
 OS Streptococcus mutans.
 XX
 PN US2005031633-A1.
 XX
 PD 10-FEB-2005.
 XX

```

PF 09-MAR-2004; 2004US-00797821.
PR 13-APR-1998; 98US-0081550P.
PR 08-JAN-1999; 99US-0115142P.
PR 12-APR-1999; 99US-00290049.
PR 07-MAR-2002; 2002US-0363209P.
PR 08-AUG-2002; 2002US-0402483P.
PR 07-MAR-2003; 2003US-00383930.
XX (SMIT/ ) SMITH D J.
PA (TAUBS/ ) TAUBMAN M A.
XX Smith DJ, Taubman MA;
XX WPI; 2005-151644/16.
XX
XX New composition comprising a fragment of a glucan binding protein-B
PT (GbpB) that binds to MHC class II protein, and a biocompatible
PT microparticle, useful for producing an antibody (claimed) for immunizing
PT mammals against dental caries.
XX
XX Claim 3; SEQ ID NO 33; 73pp; English.
XX
XX The invention relates to a composition comprising a fragment of a glucan
CC binding protein-B (GbpB) and a biocompatible microparticle, where the
CC fragment binds to a major histocompatibility complex (MHC) class II
CC protein. The composition is useful for producing an antibody for
CC immunizing mammals against dental caries. This sequence corresponds to a
CC Streptococcus mutans GbpB protein of the invention.
XX
XX Sequence 431 AA;
SQ
Query Match 100.0%; Score 92; DB 9; Length 431;
Best Local Similarity 100.0%; Pred. No. 1e-06;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 LSSATTLTSAVKADDFDAQIA 20
DB 16 LSSATTLTSAVKADDFDAQIA 35
RESULT 9
AEB91500
ID AEB91500 standard; protein; 431 AA.
XX
XX AEB91500;
XX
XX 20-OCT-2005 (first entry)
XX
XX Microbial pathogen adhesin protein sequence, SEQ ID NO:210.
XX
XX algorithm; adhesin; pharmaceutical; vaccine; drug screening;
KW bordetella pertussis infection; antibacterial; pneumonia;
KW antiinflammatory; respiratory-gen.; gastric ulcer; antitumor;
KW gastrointestinal-gen.; urinary tract infection; antimicrobial; uropathic.
XX
XX Streptococcus mutans.
XX
XX WO2005076010-A2.
XX
XX 18-AUG-2005.
XX
XX 07-FEB-2005; 2005WO-IN000037.
XX
XX 06-FEB-2004; 2004IN-DE000173.
XX 20-JUL-2004; 2004US-0589227P.
XX
XX (COUL ) COUNCIL SCI & IND RES SOUTH AFRICA.
XX
XX Sachdeva G, Kumar K, Jain P, Brahmachari SK, Ramachandran S;
XX WPI; 2005-597835/61.
XX

```

```

PT Computational method for identifying adhesin and adhesin like molecules,
PT comprises computing sequence-based attributes of protein sequences using
PT neural network software and training an artificial neural network.
XX
XX Claim 16; SEQ ID NO 210; 402pp; English.
XX
XX The present invention relates to a computational method (M1) for
CC identifying adhesin and adhesin-like proteins, by computing the sequence-
CC based attributes of protein sequences using five attribute modules of a
CC neural network software, training an artificial neural network (ANN) for
CC each of the computed five attributes, and identifying the adhesin and
CC adhesin-like proteins having probability of being an adhesin (Pad) as
CC equal or greater than 0.51. Also claimed is a set of 274 annotated genes
CC encoding adhesin and adhesin-like proteins, having 274 fully defined 162-
CC 11151 base pair (SEQ ID NO: 385-658) sequences; a set of 105 hypothetical
CC genes encoding adhesin and adhesin-like proteins, having 105 fully
CC defined 306-15876 base pairs (SEQ ID NO: 659-763); a set of 279 annotated
CC adhesin and adhesin-like proteins, having 279 fully defined 53-3716 base
CC pair (SEQ ID NO: 1-279); a set of 105 hypothetical adhesin and adhesin-
CC like proteins, having 105 fully defined 106-5291 base pair (SEQ ID NO:
CC 280-384) sequences; and a fully connected multilayer feed forward ANN (1)
CC based on (M1). (M1) is useful for identifying adhesin and adhesin-like
CC proteins, of therapeutic potential, and identifying and short-listing
CC to eliminate diseases caused by various pathogenic organisms. (M1) is
CC useful for identifying putative adhesins that are important in drug
CC discovery and preventing therapeutics for whooping cough, pneumonia,
CC gastric ulcer and urinary tract infections. (M1) identifies adhesins from
CC distantly related organisms, and from bacteria belonging to a wide
CC phylogenetic spectrum. (M1) is capable of predicting adhesive nature of
CC unique proteins. The present sequence is a microbial pathogen adhesin
CC protein sequence.
XX
XX Sequence 431 AA;
SQ
Query Match 100.0%; Score 92; DB 9; Length 431;
Best Local Similarity 100.0%; Pred. No. 1e-06;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 LSSATTLTSAVKADDFDAQIA 20
DB 16 LSSATTLTSAVKADDFDAQIA 35
RESULT 10
ADD93652
ID ADD93652 standard; protein; 432 AA.
XX
XX ADD93652;
XX
XX 29-JAN-2004 (first entry)
XX
XX Streptococcus mutans glucan binding protein-B.
XX
XX Glucan binding protein-B; GbpB; vaccine; anticaries; epitope; immunogen.
XX
XX Streptococcus mutans.
XX
XX WO2003075845-A2.
XX
XX 18-SEP-2003.
XX
XX 07-MAR-2003; 2003WO-US006962.
XX
XX 07-MAR-2002; 2002US-0363209P.
XX 08-AUG-2002; 2002US-0402483P.
XX
XX (FORS-) FORSYTH INST.
XX
XX Smith DJ, Taubman MA;
XX
XX WPI; 2003-845091/78.
XX GENBANK; AY046413.
XX

```

XX Composition useful as vaccines for dental caries comprises a fragment of
PT a glucan binding protein-B binding to a major histocompatibility complex
PT class II protein.
XX
XX Claim 5; Page 8; 49pp; English.
XX
CC The present sequence is the protein sequence of the glucan binding
CC protein-B (GbpB) of Streptococcus mutans strain 3SNI. The sequence
CC includes a number of human leucocyte antigen (HLA)-binding epitopes. The
CC invention provides immunogenic compositions and vaccines for dental
CC caries. The compositions comprise major histocompatibility complex (MHC)
CC class II protein-binding GbpB peptides, especially HLA-binding peptides,
CC covalently linked with peptide subunits of a glucosyltransferase. The
CC compositions are used in a claimed method of eliciting production of an
CC antibody in a mammal. Dieptopic or multiepitopic polypeptides can be
CC prepared synthetically or by recombinant DNA technology. Antibodies
CC raised against MHC class II binding fragments of GbpB can be used in
CC passive immunisation.
XX
XX Sequence 432 AA;
SQ
Query Match 100.0%; Score 92; DB 7; Length 432;
Best Local Similarity 100.0%; Pred. No. 1e-06;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 LSSATTL SAVKADDFDAQIA 20
DB 16 LSSATTL SAVKADDFDAQIA 35
|||||
RESULT 11
ADX37275
ID ADX37275 standard; protein; 432 AA.
XX
AC ADX37275;
XX
XX 21-APR-2005 (first entry)
XX
DE Streptococcus mutant glucan binding protein B variant #4.
XX
XX immunogenicity; immune stimulation; glucan binding protein-B;
KW microparticle; major histocompatibility complex; tooth disease.
XX
XX Streptococcus mutans.
XX
XX US2005031633-A1.
XX
XX 10-FEB-2005.
XX
XX 09-MAR-2004; 2004US-00797821.
XX
XX 13-APR-1998; 98US-0081550P.
XX
XX 08-JAN-1999; 99US-0115142P.
XX
XX 12-APR-1999; 99US-00290049.
XX
XX 07-MAR-2002; 2002US-0363209P.
XX
XX 08-AUG-2002; 2002US-0402483P.
XX
XX 07-MAR-2003; 2003US-00383930.
XX
XX (SMIT/) SMITH D J.
XX (TAUB/) TAUBMAN M A.
XX
XX Smith DJ, Taubman MA;
XX
XX WPI; 2005-151644/16.
XX
XX New composition comprising a fragment of a glucan binding protein-B
PT (GbpB) that binds to MHC class II protein, and a biocompatible
PT microparticle, useful for producing an antibody (claimed) for immunizing
PT mammals against dental caries.
XX
XX Claim 3; SEQ ID NO 32; 73pp; English.
XX

CC The invention relates to a composition comprising a fragment of a glucan
CC binding protein-B (GbpB) and a biocompatible microparticle, where the
CC fragment binds to a major histocompatibility complex (MHC) class II
CC protein. The composition is useful for producing an antibody for
CC immunizing mammals against dental caries. This sequence corresponds to a
CC Streptococcus mutans GbpB protein of the invention.
XX
XX Sequence 432 AA;
SQ
Query Match 100.0%; Score 92; DB 9; Length 432;
Best Local Similarity 100.0%; Pred. No. 1e-06;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 LSSATTL SAVKADDFDAQIA 20
DB 16 LSSATTL SAVKADDFDAQIA 35
|||||
RESULT 12
ADD93651
ID ADD93651 standard; protein; 432 AA.
XX
AC ADD93651;
XX
XX 29-JAN-2004 (first entry)
XX
XX Streptococcus mutans glucan binding protein-B.
DE
XX Glucan binding protein-B; GbpB; vaccine; anticaries; epitope; immunogen.
KW
XX Streptococcus mutans.
XX
XX WO2003075845-A2.
XX
XX 18-SEP-2003.
XX
XX 07-MAR-2003; 2003WO-US006962.
XX
XX 07-MAR-2002; 2002US-0363209P.
XX
XX 08-AUG-2002; 2002US-0402483P.
XX
XX (FORS-) FORSYTH INST.
XX
XX Smith DJ, Taubman MA;
XX
XX WPI; 2003-845091/78.
XX GENE BANK; AY046412.
XX
XX Composition useful as vaccines for dental caries comprises a fragment of
PT a glucan binding protein-B binding to a major histocompatibility complex
PT class II protein.
XX
XX Claim 5; Page 8; 49pp; English.
XX
XX The present sequence is the protein sequence of the glucan binding
CC protein-B (GbpB) of Streptococcus mutans strain 15JP2. The sequence
CC includes a number of human leucocyte antigen (HLA)-binding epitopes. The
CC invention provides immunogenic compositions and vaccines for dental
CC caries. The compositions comprise major histocompatibility complex (MHC)
CC class II protein-binding GbpB peptides, especially HLA-binding peptides,
CC covalently linked with peptide subunits of a glucosyltransferase. The
CC compositions are used in a claimed method of eliciting production of an
CC antibody in a mammal. Dieptopic or multiepitopic polypeptides can be
CC prepared synthetically or by recombinant DNA technology. Antibodies
CC raised against MHC class II binding fragments of GbpB can be used in
CC passive immunisation.
XX
XX Sequence 432 AA;
SQ
Query Match 98.9%; Score 91; DB 7; Length 432;
Best Local Similarity 95.0%; Pred. No. 1.5e-06;
Matches 19; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 LSSATTLSAVKADDFDAQIA 20
Db 16 LSSATTLSAIAKADDFDAQIA 35

RESULT 13
ADX37274
ID ADX37274 standard; protein; 432 AA.
XX AC ADX37274;
XX DT 21-APR-2005 (first entry)
XX DE Streptococcus mutant glucan binding protein B variant #3.
XX KW immunogenicity; immune stimulation; glucan binding protein-B;
KW microparticle; major histocompatibility complex; tooth disease.
XX OS Streptococcus mutans.
XX PN US2005031633-A1.
XX PD 10-FEB-2005.
XX PF 09-MAR-2004; 2004US-00797821.
XX PR 13-APR-1998; 98US-0081550P.
PR 08-JAN-1999; 99US-0115142P.
PR 12-APR-1999; 99US-00290049.
PR 07-MAR-2002; 2002US-0363209P.
PR 08-AUG-2002; 2002US-0402483P.
PR 07-MAR-2003; 2003US-00383930.
XX (SMIT/) SMITH D J.
PA (TAUBS/) TAUBMAN M A.
XX Smith DJ, Taubman MA;
PI WPI; 2005-151644/16.
XX New composition comprising a fragment of a glucan binding protein-B
PT (GbpB) that binds to MHC class II protein, and a biocompatible
PT microparticle, useful for producing an antibody (claimed) for immunizing
PT mammals against dental caries.
XX Claim 3; SEQ ID NO 31; 73pp; English.
XX The invention relates to a composition comprising a fragment of a glucan
CC binding protein-B (GbpB) and a biocompatible microparticle, where the
CC fragment binds to a major histocompatibility complex (MHC) class II
CC protein. The composition is useful for producing an antibody for
CC immunizing mammals against dental caries. This sequence corresponds to a
CC Streptococcus mutans GbpB protein of the invention.
XX Sequence 432 AA;
SQ Query Match 98.9%; Score 91; DB 9; Length 432;
Best Local Similarity 95.0%; Pred. No. 1.5e-06;
Matches 19; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 LSSATTLSAVKADDFDAQIA 20
Db 16 LSSATTLSAIAKADDFDAQIA 35

RESULT 14
ABB70601
ID ABB70601 standard; protein; 694 AA.
XX AC ABB70601;
XX DT 26-MAR-2002 (first entry)
XX

DE Drosophila melanogaster polypeptide SEQ ID NO 38595.
XX Drosophila; developmental biology; cell signalling; insecticide;
KW pharmaceutical.
XX Drosophila melanogaster.
OS WO200171042-A2.
PN 27-SEP-2001.
XX 23-MAR-2001; 2001WO-US009231.
XX 23-MAR-2000; 2000US-0191637P.
PR 11-JUL-2000; 2000US-00614150.
XX (PEKE) PE CORP NY.
PA Venter JC, Adams M, Li PWD, Myers EW;
XX WPI; 2001-656860/75.
PI N-PSDB; ABL14704.
XX New isolated nucleic acid detection reagent for detecting 1000 or more
PT genes from Drosophila and for elucidating cell signalling and cell-cell
PT interactions.
XX Disclosure; SEQ ID NO 38595; 21pp + Sequence Listing; English.
XX The invention relates to an isolated nucleic acid detection reagent
CC capable of detecting 1000 or more genes from Drosophila. The invention is
CC useful in developmental biology and in elucidating cell signalling and
CC cell-cell interactions in higher eukaryotes for the development of
CC insecticides, therapeutics and pharmaceutical drugs. The invention
CC discloses genomic DNA sequences (ABL16176-ABL30511), expressed DNA
CC sequences (ABL01840-ABL16175) and the encoded proteins (ABB57737-
CC ABB72072). The sequence data for this patent did not form part of the
CC printed specification, but was obtained in electronic format directly
CC from WIPO at ftp.wipo.int/pub/published_pct_sequences
XX Sequence 694 AA;
SQ Query Match 50.0%; Score 46; DB 4; Length 694;
Best Local Similarity 64.3%; Pred. No. 1e+02;
Matches 9; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 5 TTLSAVKADDFDAQ 18
Db 499 TLPSPIRADDFDAQ 512

RESULT 15
ABB71609
ID ABB71609 standard; protein; 4643 AA.
XX AC ABB71609;
XX 26-MAR-2002 (first entry)
XX Drosophila melanogaster polypeptide SEQ ID NO 41619.
DE Drosophila; developmental biology; cell signalling; insecticide;
KW pharmaceutical.
XX Drosophila melanogaster.
OS WO200171042-A2.
PN 27-SEP-2001.
XX 23-MAR-2001; 2001WO-US009231.
XX 23-MAR-2000; 2000US-0191637P.

PR 11-JUL-2000; 2000US-00614150.
XX (PEXE) PE CORP NY.
XX
XX Venter JC, Adams M, Li PWD, Myers EW;
PI
XX WPI; 2001-656860/75.
DR N-PSDB; ABL1512.
XX
PT New isolated nucleic acid detection reagent for detecting 1000 or more
PT genes from Drosophila and for elucidating cell signaling and cell-cell
PT interactions.
XX
XX Disclosure; SEQ ID NO 41619; 21pp + Sequence Listing; English.
XX
XX The invention relates to an isolated nucleic acid detection reagent
CC capable of detecting 1000 or more genes from Drosophila. The invention is
CC useful in developmental biology and in elucidating cell signalling and
CC cell-cell interactions in higher eukaryotes for the development of
CC insecticides, therapeutics and pharmaceutical drugs. The invention
CC discloses genomic DNA sequences (ABL16176-ABL30511), expressed DNA
CC sequences (ABL01840-ABL16175) and the encoded proteins (ABB57737-
CC ABB72072). The sequence data for this patent did not form part of the
CC printed specification, but was obtained in electronic format directly
CC from WIPO at ftp.wipo.int/pub/published_pct_sequences
XX
SQ Sequence 4643 AA;
Query Match 50.0%; Score 46; DB 4; Length 4643;
Best Local Similarity 61.1%; Pred. No. 1.1e+03;
Matches 11; Conservative 1; Mismatches 6; Indels 0; Gaps 0;
QY 1 LSSATTLSAVKADDFDAQ 18
Db : ||| ||| |||
3096 ISIGTTLVEVKAIDDFDQ 3113
Search completed: January 25, 2006, 19:01:43
Job time : 69.2 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2006 Compugen Ltd.

OM protein - protein search, using sw model

Run on: January 25, 2006, 18:48:55 ; Search time 9.1 Seconds
(without alignments)
211.465 Million cell updates/sec

Title: US-10-797-821-5
Perfect score: 92
Sequence: 1 LSSATLSAVKADDFDAQIA 20
Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283416 seqs, 96216763 residues
Total number of hits satisfying chosen parameters: 283416
Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries
Database : PIR 80:*
1: pir1:*
2: pir2:*
3: pir3:*
4: pir4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	47	51.1	428	2 F85485	flavoprotein, elec
2	47	51.1	428	2 F90634	flavoprotein limpo
3	47	51.1	428	2 C64725	fixC protein + Esc
4	45	48.9	428	2 AG0511	FixC protein limpo
5	43	46.7	398	2 E87403	hypothetical prote
6	43	46.7	1196	2 A29130	beta-amylase (EC 3
7	42	45.7	427	2 H90728	probable pectinest
8	42	45.7	427	2 A85580	probable pectinest
9	42	45.7	427	2 D64813	ybhC protein precu
10	42	45.7	837	2 H87638	TonB-dependent rec
11	42	45.7	840	2 A87639	TonB-dependent rec
12	42	45.7	898	2 A40114	fasciclin II precu
13	42	45.7	2422	2 T12687	ALR protein homolo
14	41	44.6	167	2 A11167	hypothetical prote
15	41	44.6	167	2 AD1526	hypothetical prote
16	41	44.6	310	1 S37695	calcium-binding pr
17	41	44.6	507	2 S05542	hypothetical prote
18	41	44.6	883	2 T37208	hypothetical prote
19	41	44.6	2514	1 MNWV2	nonstructural poly
20	40	43.5	198	2 S60923	hypothetical prote
21	40	43.5	218	1 B41316	flagellin B1 precu
22	40	43.5	254	2 T15938	hypothetical prote
23	40	43.5	290	2 F84839	hypothetical prote
24	40	43.5	611	2 A53418	calmegin precursor
25	40	43.5	815	2 T00538	probable serine pr
26	40	43.5	816	2 T08978	serine proteinase
27	40	43.5	2492	1 A44213	nonstructural poly
28	40	43.5	2492	1 C44213	nonstructural poly
29	40	43.5	2492	1 MNWVTD	nonstructural poly

30	40	43.5	2493	2 S72349	nonstructural poly
31	40	43.5	2493	2 S26372	nonstructural poly
32	40	43.5	2512	1 MNWVS	nonstructural poly
33	40	43.5	2514	1 MNWV82	nonstructural poly
34	39	42.4	86	2 E95225	hypothetical prote
35	39	42.4	86	2 H98089	hypothetical prote
36	39	42.4	114	2 S77813	probable ribonucle
37	39	42.4	153	2 S75357	hypothetical prote
38	39	42.4	186	2 H98220	probable proteinas
39	39	42.4	186	2 AH3025	proteinase [import
40	39	42.4	297	2 F64750	regulatory protein
41	39	42.4	329	2 AG3055	succinoglycan bios
42	39	42.4	337	2 F98230	succinoglycan bios
43	39	42.4	341	2 T41286	malate dehydrogena
44	39	42.4	621	2 JQ1685	anthranilate synth
45	39	42.4	621	2 S27752	anthranilate synth

ALIGNMENTS

RESULT 1

F85485
Flavoprotein, electron transport [imported] - Escherichia coli (strain O157:H7, substra
C:Species: Escherichia coli
C>Date: 16-Feb-2001 #sequence_revision 16-Feb-2001 #text_change 09-Jul-2004
C:Accession: F85485
R:Perna, N.T.; Plunkett III, G.; Burland, V.; Mau, B.; Glasner, J.D.; Rose, D.J.; Mayhe
iller, L.; Grotbeck, E.J.; Davis, N.W.; Lim, A.; Dimalanta, E.; Potamousis, K.; Apodaca,
Nature 409, 529-533, 2001
A:Title: Genome sequence of enterohemorrhagic Escherichia coli O157:H7.
A:Reference number: A85480; MUID:21074935; PMID:11206551
A:Accession: F85485
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-428 <STO>
A:Cross-references: UNIPROT:Q8XA26; UNIPARC:UPI00001656C3; GB:AE005174; NID:gl2512733; I
A:Experimental source: strain O157:H7, substrain EDU933
C:Genetics:
A:Gene: fixC
C:Superfamily: fixC protein

Query Match 51.1%; Score 47; DB 2; Length 428;
Best Local Similarity 52.9%; Pred. No. 4.8;
Matches 9; Conservative 4; Mismatches 4; Indels 0; Gaps 0;

QY 2 SSATLSAVKADDFDAQ 18
::|::|::|::|
Db 327 AAKTVLSAWKSDDFSKQ 343

RESULT 2

F90634
Flavoprotein [imported] - Escherichia coli (strain O157:H7, substrain RIMD 0509952)
C:Species: Escherichia coli
C>Date: 18-Jul-2001 #sequence_revision 18-Jul-2001 #text_change 09-Jul-2004
C:Accession: F90634
R:Havashi, T.; Makino, K.; Ohnishi, M.; Kurokawa, K.; Iehii, K.; Yokoyama, K.; Han, C.G
gasawara, N.; Yasunaga, T.; Kuhara, S.; Shiba, T.; Hattori, M.; Shinagawa, H.
DNA Res. 8, 11-22, 2001
A:Title: Complete genome sequence of enterohemorrhagic Escherichia coli O157:H7 and gen
A:Reference number: A99629; MUID:21156231; PMID:11258796
A:Accession: F90634
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-428 <HAY>
A:Cross-references: UNIPROT:Q8XA26; UNIPARC:UPI00000075B; GB:BA000007; PIDN:BA033469.1
A:Experimental source: strain O157:H7, substrain RIMD 0509952
C:Genetics:
A:Gene: ECs0046
C:Superfamily: fixC protein

Query Match 51.1%; Score 47; DB 2; Length 428;

```

Best Local Similarity 52.9%; Pred. No. 4.8;
Matches 9; Conservative 4; Mismatches 4; Indels 0; Gaps 0;

QY 2 SSATTL$AVKADDFDAQ 18
   ::|||::|||
Db 327 AAKTVLSAMKSDDFSKQ 343

RESULT 3
C64725
fixC protein - Escherichia coli (strain K-12)
N:Contains: probable guinone reductase (EC 1.6.5.-)
C:Species: Escherichia coli
C>Date: 12-Sep-1997 #sequence_revision 17-Sep-1997 #text_change 09-Jul-2004
C:Accession: C64725; S40564
R:Blattner, F.R.; Plunkett III, G.; Bloch, C.A.; Perna, N.T.; Burland, V.; Riley, M.; Cohen, A.; Rose, D.J.; Mau, B.; Shao, Y.
Science 277, 1453-1462, 1997
A:Title: The complete genome sequence of Escherichia coli K-12.
A:Reference number: A64720; MUID:97426617; PMID:9278503
A:Accession: C64725
A:Status: nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-428 <BLAT>
A:Cross-references: UNIPROT:P31575; UNIPARC:UPI000012A7EB; GB:AE000114; GB:U000096; NID:9
A:Experimental source: strain K-12, substrain MG1655
R:Yura, T.; Mori, H.; Nagai, H.; Nagata, T.; Ishihama, A.; Fujita, N.; Isono, K.; Mizobu
submitted to the EMBL Data Library, December 1992
A:Description: Systematic sequencing of the Escherichia coli genome: analysis of the 0-2
A:Reference number: S40531
A:Accession: S40564
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-279,'IM',282-309,'P',311-428 <YUR>
A:Cross-references: UNIPARC:UPI000016EDD3; EMBL:D10483; NID:g216434; PIDN:BAAO1319.1; PI
C:Genetics:
A:Gene: fixC
C:Superfamily: fixC protein
C:Keywords: FAD; flavoprotein; oxidoreductase; transmembrane protein
F:8-24/Domain: transmembrane #status predicted <TMM>

Query Match 51.1%; Score 47; DB 2; Length 428;
Best Local Similarity 52.9%; Pred. No. 4.8;
Matches 9; Conservative 4; Mismatches 4; Indels 0; Gaps 0;

QY 2 SSATTL$AVKADDFDAQ 18
   ::|||::|||
Db 327 AAKTVLSAMKSDDFSKQ 343

RESULT 4
AG0511
FixC protein [imported] - Salmonella enterica subsp. enterica serovar Typhi (strain CT18
C:Species: Salmonella enterica subsp. enterica serovar Typhi
A:Note: this species has also been called Salmonella typhi
C>Date: 09-Nov-2001 #sequence_revision 09-Nov-2001 #text_change 18-Nov-2002
C:Accession: AG0511
R:Parkhill, J.; Dougan, G.; James, K.D.; Thomson, N.R.; Pickard, D.; Wain, J.; Churcher,
th, T.; Connor, P.; Cronin, A.; Davis, P.; Davies, R.M.; Dowd, L.; White, N.; Farrar,
S.; Moule, S.; O'Gaora, P.
Nature 413, 848-852, 2001
A:Authors: Parry, C.; Quail, M.; Rutherford, K.; Simmonds, M.; Skelton, J.; Stevens, K.;
A:Title: Complete genome sequence of a multiple drug resistant Salmonella enterica serov
A:Reference number: AB0502; MUID:21534947; PMID:11677608
A:Accession: AG0511
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-428 <PAR>
A:Cross-references: UNIPARC:UPI00000599AE; GB:AL513382; PIDN:CAD01231.1; PID:gi16501360;
C:Genetics:
A:Gene: STY0087
C:Superfamily: fixC protein

```

```

Query Match 48.9%; Score 45; DB 2; Length 428;
Best Local Similarity 52.9%; Pred. No. 10;
Matches 9; Conservative 3; Mismatches 5; Indels 0; Gaps 0;

QY 2 SSATTL$AVKADDFDAQ 18
   ::|||::|||
Db 327 AAKTVLSAMKSDDFSKQ 343

RESULT 5
E87403
hypothetical protein CC1244 [imported] - Caulobacter crescentus
C:Species: Caulobacter crescentus
C>Date: 20-Apr-2001 #sequence_revision 20-Apr-2001 #text_change 09-Jul-2004
C:Accession: E87403
R:Nierman, W.C.; Feldblyum, T.V.; Paulsen, I.T.; Nelson, K.E.; Eisen, J.; Heidelberg, J.
B.; Laub, M.T.; DeBoy, R.T.; Dodson, R.J.; Durkin, A.S.; Gwinn, M.L.; Haft, D.H.; Kolon
n, J.; Ermolaeva, M.; White, O.; Salzberg, S.L.; Shapiro, L.; Venter, J.C.; Fraser, C.M
Proc. Natl. Acad. Sci. U.S.A. 98, 4136-4141, 2001
A:Title: Complete Genome Sequence of Caulobacter crescentus.
A:Reference number: A87249; MUID:21173698; PMID:11259647
A:Accession: E87403
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-398 <STO>
A:Cross-references: UNIPROT:Q9ABV7; UNIPARC:UPI000000C730A; GB:AE005673; NID:g13422573; F
A:Genetics:
A:Gene: CC1244

Query Match 46.7%; Score 43; DB 2; Length 398;
Best Local Similarity 45.0%; Pred. No. 21;
Matches 9; Conservative 5; Mismatches 6; Indels 0; Gaps 0;

QY 1 LSSATTL$AVKADDFDAQ 20
   ::|||::|||
Db 144 LASRRMAA$VLSGDFDAYLA 163

RESULT 6
A29130
beta-amylase (EC 3.2.1.2) / alpha-amylase (EC 3.2.1.1) precursor - Bacillus polymyxa
C:Species: Bacillus polymyxa
C>Date: 25-Oct-1987 #sequence_revision 03-Mar-1994 #text_change 09-Jul-2004
C:Accession: A29130; B29130; A32251; A29108
R:Kawazu, T.; Nakanishi, Y.; Uozumi, N.; Sasaki, T.; Yamagata, H.; Tsukagoshi, N.; Uda
J. Bacteriol. 169, 1564-1570, 1987
A:Title: Cloning and nucleotide sequence of the gene coding for enzymatically active fr
A:Reference number: A29130; MUID:87165765; PMID:2435707
A:Accession: A29130
A:Molecule type: DNA
A:Residues: 1-936 <KAW>
A:Cross-references: UNIPROT:P21543; UNIPARC:UPI000017AC9F; GB:M15817
A:Experimental source: strain 72
A:Accession: B29130
A:Molecule type: protein
A:Residues: 36-50 <KA2>
A:Cross-references: UNIPARC:UPI0000017ACA0
R:Uozumi, N.; Sakurai, K.; Sasaki, T.; Takekawa, S.; Yamagata, H.; Tsukagoshi, N.; Uda
J. Bacteriol. 171, 375-382, 1989
A:Title: A single gene directs synthesis of a precursor protein with beta- and alpha-am
A:Reference number: A32251; MUID:89123046; PMID:2464578
A:Accession: A32251
A:Status: not compared with conceptual translation
A:Molecule type: DNA
A:Residues: 937-1196 <UOZ>
A:Cross-references: UNIPARC:UPI000017ACA1
A:Experimental source: strain 72
R:Rhodes, C.; Strasser, J.; Friedberg, F.
Nucleic Acids Res. 15, 3934, 1987
A:Title: Sequence of an active fragment of B. polymyxa beta amylase.
A:Reference number: A29108; MUID:87231094; PMID:2438660
A:Accession: A29108
A:Molecule type: DNA

```

A;Residues: 'MIGL', 2-66, 'S', 68-99, 'D', 101-153, 'N', 155-176, 'Q', 178-226, 'KS', 229-329, 'S', 3
 A;Cross-References: UNIPARC:UPI0000170CFF
 A;Experimental source: ATCC 8523
 C;Genetics:
 A;Start codon: TTG
 C;Function:
 A;Description: catalyzes both the hydrolysis of internal 1,4-alpha-D-glucosidic bonds (a
 vely from the non-reducing end (beta-amyase activity)
 A;Pathway: glycogen/starch degradation
 C;Keywords: glycosidase; hydrolase; polysaccharide degradation
 F;1-35/Domain: signal sequence #status predicted <SIG>
 F;36-1196/Product: beta-amyase/alpha-amyase #status predicted <MAT>
 F;908-1033/Domain: alpha-amyase core homology <AMY>

Query Match 46.7%; Score 43; DB 2; Length 1196;
 Best Local Similarity 50.0%; Pred. No. 69;
 Matches 9; Conservative 2; Mismatches 7; Indels 0; Gaps 0;

QY 2 SSATTL5AVKADDFDAQI 19
 |||||:|||||
 Db 28 SPSNTASAAVADDFQASV 45

RESULT 7
 H90728
 probable pectinesterase [imported] - Escherichia coli (strain O157:H7, substrain RIMD 05
 C;Species: Escherichia coli
 C;Date: 18-Jul-2001 #sequence_revision 18-Jul-2001 #text_change 09-Jul-2004
 C;Accession: H90728
 R;Hayashi, T.; Makino, K.; Ohnishi, M.; Kurokawa, K.; Ishii, K.; Yokoyama, K.; Han, C.G.
 gasawara, N.; Yasunaga, T.; Kuhara, S.; Shiba, T.; Hattori, M.; Shingawa, H.
 DNA Res. 8, 11-22, 2001
 A;Title: Complete genome sequence of enterohemorrhagic Escherichia coli O157:H7 and geno
 A;Reference number: A99629; MUID:21156231; PMID:11258796
 A;Accession: H90728
 A;Status: preliminary
 A;Molecule type: DNA
 A;Residues: 1-427 <HAY>
 A;Cross-References: UNIPROT:Q8X891; UNIPARC:UPI00000D097C; GB:BA000007; PIDN:BAB34223.1;
 A;Experimental source: strain O157:H7, substrain RIMD 0509952
 C;Genetics:
 A;Gene: ECS0800
 C;Superfamily: pectinesterase pamb

Query Match 45.7%; Score 42; DB 2; Length 427;
 Best Local Similarity 58.8%; Pred. No. 34;
 Matches 10; Conservative 2; Mismatches 5; Indels 0; Gaps 0;

QY 2 SSATTL5AVKADDFDAQ 18
 |||||:|||||
 Db 40 SSRPILSAKEAQNFDQA 56

RESULT 8
 A85580
 probable pectinesterase ybhc [imported] - Escherichia coli (strain O157:H7, substrain ED
 C;Species: Escherichia coli
 C;Date: 16-Feb-2001 #sequence_revision 16-Feb-2001 #text_change 09-Jul-2004
 C;Accession: A85580
 R;Perna, N.T.; Plunkett III, G.; Burland, V.; Mau, B.; Glasner, J.D.; Rowe, D.J.; Mayhew
 iller, L.; Grotbeck, E.J.; Davis, N.W.; Lim, A.; Dimalanta, E.; Potamousis, K.; Apodaca,
 Nature 409, 529-533, 2001
 A;Title: Genome sequence of enterohemorrhagic Escherichia coli O157:H7.
 A;Reference number: A85480; MUID:21074935; PMID:11206551
 A;Accession: A85580
 A;Status: preliminary
 A;Molecule type: DNA
 A;Residues: 1-427 <STO>
 A;Cross-References: UNIPROT:Q8X891; UNIPARC:UPI00000D097C; GB:AE005174; NID:gl2513707; F
 A;Experimental source: strain O157:H7, substrain EDL933
 C;Genetics:
 A;Gene: ybhc

C;Superfamily: pectinesterase pamb

Query Match 45.7%; Score 42; DB 2; Length 427;
 Best Local Similarity 58.8%; Pred. No. 34;
 Matches 10; Conservative 2; Mismatches 5; Indels 0; Gaps 0;

QY 2 SSATTL5AVKADDFDAQ 18
 |||||:|||||
 Db 40 SSRPILSAKEAQNFDQA 56

RESULT 9

D64813

ybhc protein precursor - Escherichia coli (strain K-12)

C;Species: Escherichia coli

C;Date: 12-Sep-1997 #sequence_revision 17-Sep-1997 #text_change 09-Jul-2004

C;Accession: D64813

R;Blattner, F.R.; Plunkett III, G.; Bloch, C.A.; Perna, N.T.; Burland, V.; Riley, M.; C

A.; Rose, D.J.; Mau, B.; Shao, Y.

Science 277, 1453-1462, 1997

A;Title: The complete genome sequence of Escherichia coli K-12.

A;Reference number: A64720; MUID:97426617; PMID:9278503

A;Accession: D64813

A;Status: nucleic acid sequence not shown; translation not shown

A;Molecule type: DNA

A;Residues: 1-427 <BLAT>

A;Cross-References: UNIPROT:P46130; UNIPARC:UPI000013A348; GB:AE000180; NID:

A;Experimental source: strain K-12, substrain MG1655

C;Genetics:

A;Gene: ybhc

C;Superfamily: pectinesterase pamb

F;1-22/Domain: signal sequence #status predicted <SIG>

F;23-427/Product: ybhc protein #status predicted <MAT>

Query Match 45.7%; Score 42; DB 2; Length 427;
 Best Local Similarity 58.8%; Pred. No. 34;
 Matches 10; Conservative 2; Mismatches 5; Indels 0; Gaps 0;

QY 2 SSATTL5AVKADDFDAQ 18
 |||||:|||||
 Db 40 SSRPILSAKEAQNFDQA 56

RESULT 10

H87638

TonB-dependent receptor [imported] - Caulobacter crescentus

C;Species: Caulobacter crescentus

C;Date: 20-Apr-2001 #sequence_revision 20-Apr-2001 #text_change 09-Jul-2004

C;Accession: H87638

R;Nierman, W.C.; Feldblyum, T.V.; Paulsen, I.T.; Nelson, K.E.; Eisen, J.; Heidelberg, J.

B.; Laub, M.T.; DeBoy, R.T.; Dodson, R.J.; Durkin, A.S.; Gwinn, M.L.; Haft, D.H.; Kolo

n, J.; Ermolaeva, M.; White, O.; Salzberg, S.L.; Shapiro, L.; Venter, J.C.; Fraser, C.M

Proc. Natl. Acad. Sci. U.S.A. 98, 4136-4141, 2001

A;Title: Complete Genome Sequence of Caulobacter crescentus.

A;Reference number: A87249; MUID:21173698; PMID:11259647

A;Accession: H87638

A;Status: preliminary

A;Molecule type: DNA

A;Residues: 1-837 <STO>

A;Cross-References: UNIPROT:Q9A3Q6; UNIPARC:UPI00000C798B; GB:AE005673; NID:gl3424812;

C;Genetics:

A;Gene: CC3146

Query Match 45.7%; Score 42; DB 2; Length 837;
 Best Local Similarity 50.0%; Pred. No. 70;
 Matches 8; Conservative 5; Mismatches 3; Indels 0; Gaps 0;

QY 1 LSSATTL5AVKADDFD 16
 |||||:|||||
 Db 60 LDSATSTGALRASEFE 75

RESULT 11

C:Genetics:
A:Gene: lin0748

Query Match 44.6%; Score 41; DB 2; Length 167;
Best Local Similarity 66.7%; Pred. No. 18;
Matches 8; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 5 TTLSAVKADDFD 16
| | | | | | | | | |
Db 83 TKLDKVKADDYD 94

Search completed: January 25, 2006, 19:15:17
Job time : 10.1 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2006 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: January 25, 2006, 18:48:24 ; Search time 70.3 Seconds
(without alignments)
200.719 Million cell updates/sec

Title: US-10-797-821-5

Perfect score: 92

Sequence: 1 LSSATTL SAVKADDFDAQIA 20

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 2166443 seqs, 705528306 residues

Total number of hits satisfying chosen parameters: 2166443

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : UniProt_05.80.*

1: uniprot_sprot.*

2: uniprot_trembl.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	92	100.0	211	2 Q9ZAS7_STRMU	Q9ZAS7 streptococc
2	92	100.0	431	2 Q938V0_STRMU	Q938V0 streptococc
3	92	100.0	431	2 Q938V3_STRMU	Q938V3 streptococc
4	92	100.0	431	2 Q9AG98_STRMU	Q9AG98 streptococc
5	92	100.0	431	2 Q8DWM3_STRMU	Q8DWM3 streptococc
6	92	100.0	432	2 Q938V1_STRMU	Q938V1 streptococc
7	91	98.9	432	2 Q938V2_STRMU	Q938V2 streptococc
8	50	54.3	187	2 Q6NA84_RHOPA	Q6NA84 rhodospaudo
9	49.5	53.8	474	2 Q5M6K4_STRT2	Q5M6K4 streptococc
10	49.5	53.8	485	2 Q5M212_STRT1	Q5M212 streptococc
11	48	52.2	462	2 Q4P3F1_USTMA	Q4P3F1 ustilago ma
12	47	51.1	324	2 Q4Q099_LEIMA	Q4Q099 leishmania
13	47	51.1	428	1 FIXC_EGOS7	Q7ant0 escherichia
14	47	51.1	428	1 FIXC_ECOL6	P68645 escherichia
15	47	51.1	428	1 FIXC_ECOLI	P68644 escherichia
16	47	51.1	428	1 FIXC_SHIFL	Q83aq7 shigella fl
17	47	51.1	1449	1 DP03_CLOPE	Q8xir3 clostridium
18	46	50.0	421	2 Q9SU86_DROME	Q9SU86 drosophila
19	46	50.0	452	2 Q88YJ2_LACPL	Q88YJ2 lactobacill
20	46	50.0	675	2 Q4WPR6_ASEFU	Q4WPR6 aspergillus
21	46	50.0	719	2 Q9U4G1_DROSOPH	Q9U4G1 drosophila
22	46	50.0	814	1 SYEB_SYNPX	Q7u6v9 synechococc
23	46	50.0	4705	1 FAT2_DROME	Q9wv71 drosophila
24	45	48.9	253	2 QSH3X0_XANOR	Q5h3x0 xanthomonas
25	45	48.9	428	1 FIXC_SALTI	Q829k9 salmonella
26	45	48.9	428	1 FIXC_SALTY	Q829w9 salmonella
27	45	48.9	428	2 Q57Ti3_SALCH	Q57Ti3 salmonella
28	45	48.9	428	2 Q5PIN9_SALPA	Q5pin9 salmonella
29	45	48.9	447	2 Q9AKA4_STRAG	Q9aka4 streptococc
30	45	48.9	447	2 Q8E2H1_STRAS	Q8e2h1 streptococc
31	45	48.9	447	2 Q8E7X9_STRAS3	Q8e7x9 streptococc

32	45	48.9	516	2 Q7QFZ8_ANOGA	Q7qfz8 anopheles g
33	45	48.9	1731	2 Q4QJ23_LEIMA	Q4qj23 leishmania
34	44	47.8	149	2 Q5E103_GECJA	Q5e103 gecko japon
35	44	47.8	233	2 Q4NBV5_9MICC	Q4nbv5 arthrobacte
36	44	47.8	427	1 PROA_GLUOX	Q5ftr2 gluconobact
37	44	47.8	446	2 Q7VW07_BORPE	Q7vw07 bordetella
38	44	47.8	459	2 Q7W544_BORPA	Q7w544 bordetella
39	44	47.8	459	2 Q7WCM6_BORBR	Q7wcm6 bordetella
40	44	47.8	718	2 Q521A4_MAGGR	Q521a4 magnaporthe
41	44	47.8	1589	2 Q4FXY1_LEIMA	Q4fxy1 leishmania
42	44	47.8	2053	2 Q4QDU5_LEIMA	Q4qdu5 leishmania
43	43	46.7	153	2 Q87UB6_PSESM	Q87ub6 pseudomonas
44	43	46.7	186	2 Q6ANP0_DSPPS	Q6anp0 desulfotale
45	43	46.7	272	2 Q53LB0_ORYSA	Q53lb0 oryza sativ

ALIGNMENTS

RESULT 1
Q9ZAS7_STRMU ID Q9ZAS7_STRMU PRELIMINARY; PRT; 211 AA.
AC Q9ZAS7;
DT 01-MAY-1999 (Tremblrel. 10, Created)
DT 01-MAY-1999 (Tremblrel. 10, Last sequence update)
DT 01-OCT-2002 (Tremblrel. 22, Last annotation update)
DE Putative secreted protein (fragment).
OS Streptococcus mutans.
OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
OC Streptococcus.
OX NCBI_TaxID=1309;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=GS-5 Kuramitsu;
RA Peruzzi F., Piggot P.J., Daneo-Moore L.;
RL Submitted (NOV-1996) to the EMBL/GenBank/DBJ databases.
DR EMBL; U78607; AAD00288.1; -; Genomic_DNA.
FT NON_TER 211
SQ SEQUENCE 211 AA; 22803 MW; 4ACE331159CFAPC6 CRC64;

Query Match 100.0%; Score 92; DB 2; Length 211;
Best Local Similarity 100.0%; Pred. No. 7.6e-07;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 LSSATTL SAVKADDFDAQIA 20
|||||
Db 16 LSSATTL SAVKADDFDAQIA 35

RESULT 2
Q938V0_STRMU ID Q938V0_STRMU PRELIMINARY; PRT; 431 AA.
AC Q938V0;
DT 01-DEC-2001 (Tremblrel. 19, Created)
DT 01-DEC-2001 (Tremblrel. 19, Last sequence update)
DT 01-WAR-2004 (Tremblrel. 26, Last annotation update)
DE Glucan-binding protein B.
OS Streptococcus mutans.
OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
OC Streptococcus.
OX NCBI_TaxID=1309;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=5SM3;
RA Jin S., Duncan M.J., Taubman M.A., Smith D.J.;
RT "Cloning of the gbpB gene from Streptococcus mutans.";
RL J. Dent. Res. 79:224-224 (2000).
RN [2]
RC NUCLEOTIDE SEQUENCE.
RC STRAIN=5SM3;
RX MEDLINE=21481971; PubMed=11598068;
RX DOI=10.1128/JAI.69.11.6931-6941.2001;
RA Mattos-Graner R.O., Jin S., King W.F., Chen T., Smith D.J.,

RA Duncan M.J.;
RT "Cloning of the Streptococcus mutans gene encoding glucan binding
RT protein B and analysis of genetic diversity and protein production in
RT clinical isolates.";
RL Infect. Immun. 69:6931-6941(2001).
DR EMBL; AY046414; AAK94504.1; -; Genomic_DNA.
DR InterPro; IPR007921; CHAP.
DR InterPro; IPR009148; Siba.
DR Pfam; PF05257; CHAP; 1.
DR PRINTS; PR01852; SIBAPROTEIN.
DR PROSITE; PS050911; CHAP; 1.
SQ SEQUENCE 431 AA; 44650 MW; 05D38D8B8BC4609F CRC64;

Query Match 100.0%; Score 92; DB 2; Length 431;
Best Local Similarity 100.0%; Pred. No. 1.6e-06;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 LSSATTL\$AVKADDFDAQIA 20
DB 16 LSSATTL\$AVKADDFDAQIA 35

RESULT 3
Q938V3 STRMU PRELIMINARY; PRT; 431 AA.
ID Q938V3;
AC Q938V3;
DT 01-DEC-2001 (TrEMBLrel. 19, Created)
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DE 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Glucan-binding protein B.
OS Streptococcus mutans.
OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
OC Streptococcus.
OX NCBI_TaxID=1309;
RN NUCLEOTIDE SEQUENCE.
RC STRAIN=GS-5;
RA Jin S., Duncan M.J., Taubman M.A., Smith D.J.;
RA "Cloning of the gbpB gene from Streptococcus mutans.";
RL J. Dent. Res. 79:224-224(2000).
RN [2]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=GS-5;
RX MEDLINE=21481971; PubMed=11598068;
RX DOI=10.1128/IAI.69.11.6931-6941.2001;
RX "Cloning of the gbpB gene from Streptococcus mutans.";
RA Mattos-Graner R.O., Jin S., King W.F., Chen T., Smith D.J.,
RA Duncan M.J.;
RA "Cloning of the Streptococcus mutans gene encoding glucan binding
RT protein B and analysis of genetic diversity and protein production in
RT clinical isolates.";
RL Infect. Immun. 69:6931-6941(2001).
DR EMBL; AY046410; AAK94500.1; -; Genomic_DNA.
DR InterPro; IPR007921; CHAP.
DR InterPro; IPR009148; Siba.
DR Pfam; PF05257; CHAP; 1.
DR PRINTS; PR01852; SIBAPROTEIN.
DR PROSITE; PS050911; CHAP; 1.
SQ SEQUENCE 431 AA; 44620 MW; 464FE3B563FB7E51 CRC64;

Query Match 100.0%; Score 92; DB 2; Length 431;
Best Local Similarity 100.0%; Pred. No. 1.6e-06;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 LSSATTL\$AVKADDFDAQIA 20
DB 16 LSSATTL\$AVKADDFDAQIA 35

RESULT 4
Q9AG98 STRMU PRELIMINARY; PRT; 431 AA.
ID Q9AG98;
AC Q9AG98;
DT 01-JUN-2001 (TrEMBLrel. 17, Created)

DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
DT 01-FEB-2005 (TrEMBLrel. 29, Last annotation update)
DE Immunodominant glycoprotein IDG-60 (Glucan-binding protein B).
GN Name=saga;
OS Streptococcus mutans.
OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
OC Streptococcus.
OX NCBI_TaxID=1309;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=GS-5;
RX MEDLINE=21153617; PubMed=11254612;
RX DOI=10.1128/IAI.69.4.2493-2501.2001;
RX Chia J.S., Lee Y.Y., Huang P.T., Chen J.Y.;
RA "Identification of stress-responsive genes in Streptococcus mutans by
RT differential display reverse transcription-PCR.";
RL Infect. Immun. 69:2493-2501(2001).
RN [2]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=GS-5;
RX MEDLINE=21481977; PubMed=11598074;
RX DOI=10.1128/IAI.69.11.6987-6998.2001;
RX Chia J.S., Chang L.Y., Shun C.T., Chang Y.Y., Chen J.Y.;
RA "A 60-kilodalton immunodominant glycoprotein is essential for cell
RT wall integrity and the maintenance of cell shape in Streptococcus
RT mutans.";
RL Infect. Immun. 69:6987-6998(2001).
RN [3]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=GS-5;
RX MEDLINE=21481971; PubMed=11598068;
RX DOI=10.1128/IAI.69.11.6931-6941.2001;
RX Mattos-Graner R.O., Jin S., King W.F., Chen T., Smith D.J.,
RA Duncan M.J.;
RA "Cloning of the Streptococcus mutans gene encoding glucan binding
RT protein B and analysis of genetic diversity and protein production in
RT clinical isolates.";
RL Infect. Immun. 69:6931-6941(2001).
DR EMBL; AF338445; AAK08104.1; -; Genomic_DNA.
DR EMBL; AY046411; AAK94501.1; -; Genomic_DNA.
DR InterPro; IPR007921; CHAP.
DR InterPro; IPR009148; Siba.
DR Pfam; PF05257; CHAP; 1.
DR PRINTS; PR01852; SIBAPROTEIN.
DR PROSITE; PS050911; CHAP; 1.
SQ SEQUENCE 431 AA; 44592 MW; 3EBE21FCSB47232E CRC64;

Query Match 100.0%; Score 92; DB 2; Length 431;
Best Local Similarity 100.0%; Pred. No. 1.6e-06;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 LSSATTL\$AVKADDFDAQIA 20
DB 16 LSSATTL\$AVKADDFDAQIA 35

RESULT 5
Q8DWM3 STRMU PRELIMINARY; PRT; 431 AA.
ID Q8DWM3;
AC Q8DWM3;
DT 01-MAR-2003 (TrEMBLrel. 23, Created)
DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)

DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Putative secreted antigen GbpB/Saga; putative peptidoglycan
DE Hydrolase.
GN Name=gbpB; OrderedLocusNames=SMU.22;
OS Streptococcus mutans.
OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
OC Streptococcus.
OX NCBI_TaxID=1309;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=UA159 / ATCC 700610 / Serotype c;
RX MEDLINE=22295063; PubMed=12397186; DOI=10.1073/pnas.172501299;
RA Ajdic D.J., McShan W.M., McLaughlin R.E., Savic G., Chang J.,
RA Carson M.B., Primeaux C., Fian R., Kenton S., Jia H.G., Lin S.P.,
RA Qian Y., Li S., Zhu H., Najjar F.Z., Lai H., White J., Roe B.A.,
RA Ferretti J.J.;
RT "Genome sequence of Streptococcus mutans UA159, a cariogenic dental
RT pathogen.";
RL Proc. Natl. Acad. Sci. U.S.A. 99:14434-14439 (2002).
DR EMBL; AB014855; AAN57811.1; -; Genomic_DNA.
DR InterPro; IPR007921; CHAP.
DR InterPro; IPR009148; Siba.
DR Pfam; PF05257; CHAP; 1.
DR PRINTS; PR01852; SIBAPROTEIN.
DR PROSITE; PS50911; CHAP; 1.
KW Complete proteome.
SQ SEQUENCE 431 AA; 44620 MW; 2D1CA685248CCD3E CRC64;
Query Match 100.0%; Score 92; DB 2; Length 431;
Best Local Similarity 100.0%; Pred. No. 1.6e-06;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 LSSATTL SAVKADDFDAQIA 20
DB 16 LSSATTL SAVKADDFDAQIA 35
RESULT 6
Q938V1_STRMU PRELIMINARY; PRT; 432 AA.
AC Q938V1;
DT 01-DEC-2001 (TrEMBLrel. 19, Created)
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Glucan-binding protein B.
OS Streptococcus mutans.
OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
OC Streptococcus.
OX NCBI_TaxID=1309;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=3SN1;
RA Jin S., Duncan M.J., Taubman M.A., Smith D.J.;
RT "Cloning of the gbpB gene from Streptococcus mutans.";
RL J. Dent. Res. 79:224-224 (2000).
RN [2]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=3SN1;
RX MEDLINE=21481971; PubMed=11598068;
RX DOI=10.1128/IAI.69.11.6931-6941.2001;
RA Mattos-Graner R.O., Jin S., King W.F., Chen T., Smith D.J.,
RA Duncan M.J.;
RT "Cloning of the Streptococcus mutans gene encoding glucan binding
RT protein B and analysis of genetic diversity and protein production in
RT clinical isolates.";
RL Infect. Immun. 69:6931-6941 (2001).
DR EMBL; AY046412; AAK94503.1; -; Genomic_DNA.
DR InterPro; IPR007921; CHAP.
DR InterPro; IPR009148; Siba.
DR Pfam; PF05257; CHAP; 1.
DR PRINTS; PR01852; SIBAPROTEIN.
DR PROSITE; PS50911; CHAP; 1.
SQ SEQUENCE 432 AA; 44648 MW; E7G9B2504AE50E9 CRC64;
Query Match 98.9%; Score 91; DB 2; Length 432;
Best Local Similarity 95.0%; Pred. No. 2.4e-06;
Matches 19; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
QY 1 LSSATTL SAVKADDFDAQIA 20
DB 16 LSSATTL SAVKADDFDAQIA 35
RESULT 8
Q6NA84_RHOPA PRELIMINARY; PRT; 187 AA.
AC Q6NA84;
DT 05-JUL-2004 (TrEMBLrel. 27, Created)
DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE Putative intracellular protease, PfPI family.
GN Name=pfPI; OrderedLocusNames=RPA1301;
OS Rhodopseudomonas palustris.
OC Bacteria; Proteobacteria; Alphaproteobacteria; Rhizobiales;
OC Bradyrhizobiaceae; Rhodospseudomonas.
OX NCBI_TaxID=1076;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=CGA009 / ATCC BAA-98;
RX PubMed=14704707; DOI=10.1038/nbt923;

Query Match 100.0%; Score 92; DB 2; Length 432;
Best Local Similarity 100.0%; Pred. No. 1.6e-06;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 LSSATTL SAVKADDFDAQIA 20
DB 16 LSSATTL SAVKADDFDAQIA 35
RESULT 7
Q938V2_STRMU PRELIMINARY; PRT; 432 AA.
AC Q938V2;
DT 01-DEC-2001 (TrEMBLrel. 19, Created)
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Glucan-binding protein B.
OS Streptococcus mutans.
OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
OC Streptococcus.
OX NCBI_TaxID=1309;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=150F2;
RA Jin S., Duncan M.J., Taubman M.A., Smith D.J.;
RT "Cloning of the gbpB gene from Streptococcus mutans.";
RL J. Dent. Res. 79:224-224 (2000).
RN [2]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=150F2;
RX MEDLINE=21481971; PubMed=11598068;
RX DOI=10.1128/IAI.69.11.6931-6941.2001;
RA Mattos-Graner R.O., Jin S., King W.F., Chen T., Smith D.J.,
RA Duncan M.J.;
RT "Cloning of the Streptococcus mutans gene encoding glucan binding
RT protein B and analysis of genetic diversity and protein production in
RT clinical isolates.";
RL Infect. Immun. 69:6931-6941 (2001).
DR EMBL; AY046412; AAK94502.1; -; Genomic_DNA.
DR InterPro; IPR007921; CHAP.
DR InterPro; IPR009148; Siba.
DR Pfam; PF05257; CHAP; 1.
DR PRINTS; PR01852; SIBAPROTEIN.
DR PROSITE; PS50911; CHAP; 1.
SQ SEQUENCE 432 AA; 44648 MW; E7G9B2504AE50E9 CRC64;
Query Match 98.9%; Score 91; DB 2; Length 432;
Best Local Similarity 95.0%; Pred. No. 2.4e-06;
Matches 19; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
QY 1 LSSATTL SAVKADDFDAQIA 20
DB 16 LSSATTL SAVKADDFDAQIA 35
RESULT 8
Q6NA84_RHOPA PRELIMINARY; PRT; 187 AA.
AC Q6NA84;
DT 05-JUL-2004 (TrEMBLrel. 27, Created)
DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE Putative intracellular protease, PfPI family.
GN Name=pfPI; OrderedLocusNames=RPA1301;
OS Rhodopseudomonas palustris.
OC Bacteria; Proteobacteria; Alphaproteobacteria; Rhizobiales;
OC Bradyrhizobiaceae; Rhodospseudomonas.
OX NCBI_TaxID=1076;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=CGA009 / ATCC BAA-98;
RX PubMed=14704707; DOI=10.1038/nbt923;

RA	Larimer F.W., Chain P., Hauser L., Lamerdin J.E., Malfatti S., Do L.,
RA	Libson M.L., Pelletier D.A., Beatty J.T., Lang A.S., Tabita F.R.,
RA	Gibson J.L., Hanson T.B., Bobst C., Torres y Torres J.L., Peres C.,
RA	Harrison F.H., Gibson J., Harwood C.S.;
RT	"Complete genome sequence of the metabolically versatile
RT	photosynthetic bacterium <i>Rhodospseudomonas palustris</i> .";
RL	Nat. Biotechnol. 22:55-61(2004).
DR	EMBL; BX572597; CAE26744.1; -; Genomic_DNA.
DR	HSP; O59413; 1G2I.
DR	GO; GO:0016798; F:Hydrolase activity, acting on glycosyl bonds; IEA.
DR	GO; GO:0008233; F:peptidase activity; IEA.
DR	InterPro; IPR006286; peptidase C56.
DR	InterPro; IPR002818; ThiJ/PfpI.
DR	Pfam; PF01965; DJ-1_PfpI; 1.
DR	TIGRFAMs; TIGR01382; PfpI; 1.
KW	Complete proteome; Protease.
SO	SEQUENCE 187 AA; 20279 MW; 3DAC3649ACC1257C CRC64;

Query Match 54.3%; Score 50; DB 2; Length 187;
Best Local Similarity 90.9%; Pred. No. 6.3;
Matches 10; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 7 LSAVKADDFDA 17
|||:|
Db 65 LSAVKADDYDA 75

```

RESULT 9
Q5M6K4_STRT2
ID Q5M6K4_STRT2 PRELIMINARY; PRT; 474 AA.
AC Q5M6K4;
DT 01-FEB-2005 (TrEMBLrel. 29, Created)
DT 01-FEB-2005 (TrEMBLrel. 29, Last sequence update)
DT 13-SEP-2005 (TrEMBLrel. 31, Last annotation update)
DE Glucan binding protein (Pcsb).
GN Name=pcsb; OrderedLocusNames=sttu0022;
OS Streptococcus thermophilus (strain ATCC BAA-250 / LMG 18311).

```

Streptococcus.
NCBI TaxID=264199;

LI. NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
 RP PubMed15431133; DOI=10.1038/nbt1034;
 RX Bolotin A., Quinquis B., Renault P., Sorokin A., Ehrlich S.D.,
 RA Bolotin A., Quinquis B., Renault P., Sorokin A., Ehrlich S.D.,
 RA Kulakauskas S., Lapidus A., Goltzman E., Mazur M., Pusch G.D.,
 RA Fankhauser M., Overbeek R., Kyprides N., Purnelle B., Prozzi D.,
 EA Ngui K., Masuy D., Hancy F., Bureau S., Boutry M., Delcour J.,
 RA Goffeau A., Hols P.;
 RT "Complete sequence and comparative genome analysis of the dairy
 RT bacterium *Streptococcus thermophilus*.";
 RL Nat. Biotechnol 22:1554-1558(2004).

[2]
RN
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=LMG 18311;
RA Borges F., Lavey S., Thibessard A., Fernandez A., Gintz B., Hols P.,
RA Decaris B., Leblond-Bourget N.;
RA "cse", a Chimeric and Variable Gene, Encodes an Extracellular Protein
RT Involved in Cellular Segregation in *Streptococcus thermophilus*.";
RT *J. Bacteriol.* 187:2737-2746 (2005).
RL
RL EMBL; CP000023; AAV59752.1; -- Genomic DNA.
DR EMBL; AY730643; AAW82375.1; -- Genomic DNA.
DR

INTERPRO; IP00079241; CHAP.
 Pfam; PF05257; CHAP; 1.
 PRINTS; PR01852; SIBAPROTEIN.
 PROSITE; PS0911; CHAP; 1.
 Complete proteome.
 SEQUENCE 474 AA, 48142 MW, 1BAC6A9E0A0A200D CRC64;

Query Match 53.8%; Score 49.5; DB 2; Length 474;
 Best Local Similarity 56.5%; Pred. No. 20;
 Matches 13; Conservative 3; Mismatches 4; Indels 3; Gaps 1;

```

Qy 1 LSSATWLSA---VKADDFDAQIA 20
      | | | | | | | | | |
Db 10 LVSGVTLSAAASHVAEDYDSQIA 32

RESULT 10
Q5M212_STRT1 Q5M212_STRT1 PRELIMINARY; PRT; 485 AA.
AC Q5M212;
DT 01-FEB-2005 (TrEMBLrel. 29, Created)
DT 01-FEB-2005 (TrEMBLrel. 29, Last sequence update)
DT 01-FEB-2005 (TrEMBLrel. 29, Last annotation update)
DE Glucan binding protein.
DE Name=pcsb; OrderedLocusNames=str0022;
OS Streptococcus thermophilus (strain CNRZ 1066).
OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
OC Streptococcus.
OX NCBI_taxid=299768;
[1]
RN NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
RP PubMed:15543133; DOI=10.1038/nbt1034;
RX Bolotin A., Quinquis B., Renault P., Sorokin A., Ehrlich S.D.,
RA Kulakaustas S., Lapidus A., Golsman E., Mazur M., Pusch G.D.,
RA Fonstein M., Overbeek R., Kyprides N., Fumelle B., Prozi D.,
RA Ngui K., Masuy D., Hancy P., Berteau S., Boutry M., Delcour J.,
RA Goffeau A., Hols P.;
RT "Complete sequence and comparative genome analysis of the dairy
RT bacterium Streptococcus thermophilus.";
RL Nat. Biotechnol. 22:1554-1558 (2004).
RD EMBL; CP000024; AA061641.1; -; Genomic_DNA.
DR InterPro; IPR007921; CHAP.
DR InterPro; IPR009148; Siba.
DR Pfam; PF05257; CHAP; 1.
DR PRINTS; PR01852; SIBAPROTEIN.
DR PROSITE; PS00911; CHAP; 1.
KW Complete proteome.
SQ SEQUENCE 485 AA; 49165 MW; 8665E4B1FC6CF7CA CRC64;

```

Query Match 53.8%; Score 49.5; DB 2; Length 485;
Best Local Similarity 56.5%; Pred. No. 21;
Matches 13; Conservative 3; Mismatches 4; Indels 3; Gaps 1;

Qy 1 LSSATTLSA---VKADDFDAQIA 20
Db 10 LVSGVTLSAAASVHAEDYDSQIA 32

RESULT 11	
Q4P3F1_USTMA	
ID Q4P3F1_USTMA PRELIMINARY;	PRT; 462 AA.
AC Q4P3F1	
DT 13-SEP-2005 (TREMBLrel. 31, Created)	
DT 13-SEP-2005 (TREMBLrel. 31, Last sequence update)	
DT 13-SEP-2005 (TREMBLrel. 31, Last annotation update)	
DE Hypothetical protein.	
GN ORFNames=UM05362.1;	
OS Ustilago maydis 521.	
OC Eukaryota; Fungi; Basidiomycota; Ustilaginomycetes;	
OC Ustilaginomycetidae; Ustilaginales; Ustilaginaceae; Ustilago	
OC NCBI TaxID=237631.	
OX	

[1]

KN NP NUCLEOTIDE SEQUENCE.

RC STRAIN=521;

RA Birren B., Nusbaum C., Abebe A., Abouelleil A., Adekoya E.,

RA Birren B., Allen T., Allen T., An P., Anderson M., Anderson S.,

RA Atachchi H., Ambruster J., Bachantsang P., Baldwin J., Barry A.,

RA Bayul T., Blitshteyn B., Bloom T., Blye J., Boguslavskiy L.,

RA Bowsky M., Boukgalter B., Brunache A., Butler J., Calixte N.,

RA Calvo S., Camarata J., Campo K., Chang J., Chensatsang Y., Citroen M.,

RA Collimore A., Considine T., Cook A., Cooke P., Corum B., Cuomo C.,

RA David R., Dawoe T., Degray S., Dodge S., Dooley K., Dorje P.,

RA Dorjee K., Dorris L., Duffey N., Dupes A., Elkins T., Engels R.,

RA Ericson J., Farina A., Faro S., Ferreira P., Fischer H.,

RA Fitzgerald M., Foley K., Gage D., Galagan J., Gearin G., Gnerre S.,
RA Gnirke A., Goyette A., Graham J., Grandbois E., Gyaltzen K., Hafez M.,
RA Hagopian D., Hagos B., Hall J., Hatcher B., Heller A., Higgins H.,
RA Honan T., Horn A., Houde N., Hughes L., Hulme W., Husby E., Iliev I.,
RA Jaffe D., Jones C., Kamal M., Kamat A., Kamysseleis M., Karlsson E.,
RA Kellis C., Kieu A., Kienner P., Kodira C., Kulbokas E., Labutti K.,
RA Lama D., Landers T., Leger J., Levine S., Lewis D., Lewis T.,
RA Lindblad-toh K., Liu X., Lokysang T., Lokysang Y., Lucien O.,
RA Lui A., Ma L.J., Mabbitt R., Macdonald J., Maclean C., Major J.,
RA Manning J., Marabelli R., Maru K., Matthews C., Mauceli E.,
RA McCarthy M., McDonough S., Mcghee T., Meldrim J., Menues L.,
RA Mesirov J., Mihalev A., Mihova T., Mikkelsen T., Menga V., Moru K.,
RA Mozes J., Mulrain I., Munson G., Naylor J., News C., Nguyen C.,
RA Nguyen N., Nguyen T., Nicol R., Nielsen C., Nizzari M., Norbu C.,
RA Norbu N., O'donnell P., Okoro O., O'leary S., Omotosho B.,
RA O'Neill K., Osman S., Parker S., Perrin D., Phunkhang P., Pignani B.,
RA Purcell S., Rachupka T., Ramasamy U., Rameau R., Ray V., Raymond C.,
RA Rettal R., Richardson S., Rise C., Rodriguez J., Rogers J., Rogov P.,
RA Rutman M., Schupbach R., Seaman C., Settipalli S., Sharpe T.,
RA Sheridan J., Sherpa N., Shi J., Smirnov S., Smith C., Sougnez C.,
RA Spencer B., Stalker J., Stange-thomann N., Stavropoulos S.,
RA Stenson K., Stone C., Stone S., Stubbs M., Talamas J., Tchuinga P.,
RA Tenzing P., Tesfaye S., Theodore J., Thoulutsang Y., Topham K.,
RA Towey S., Tsamla T., Tsomo N., Vallee D., Vassiliou H.,
RA Venkataraman V., Vinson J., Vo A., Wade C., Wang S., Wangchuk T.,
RA Wangdi T., Whitaker C., Wilkinson J., Wu Y., Wyman D., Yadav S.,
RA Yang S., Yang X., Yeager S., Yee E., Young G., Zainoun J., Zembek L.,
RA Zimmer A., Zody M., Zander E.;
RT "The genome sequence of *Ustilago maydis*.";
RL Submitted (FEB-2004) to the EMBL/GenBank/DBJ databases.
CC -!- CAUTION: The sequence shown here is derived from an
CC EMBL/GenBank/DBJ whole genome shotgun (WGS) entry which is
CC preliminary data.
DR EMBL; AAC0100192; EAK86611.1; -; Genomic_DNA.
KW Hypothetical protein.
SQ SEQUENCE 462 AA; 50701 MW; EB83795F497CBAC5 CRC64;

Query Match 52.2%; Score 48; DB 2; Length 462;
Best Local Similarity 53.3%; Pred. No. 35;
Matches 8; Conservative 5; Mismatches 2; Indels 0; Gaps 0;

QY 2 SSATTL5AVKADDFD 16
DB 242 SDASSLSASIKVDVDFD 256

RESULT 12
Q4Q099_LEIMA PRELIMINARY; PRT; 324 AA.
AC Q4Q099;
DT 13-SEP-2005 (TrEMBLrel. 31, Created)
DT 13-SEP-2005 (TrEMBLrel. 31, Last sequence update)
DE Hypothetical protein.
GN OrlNames-LmjF36.6560;
OS Leishmania major.
OC Eukaryota; Euklenozoa; Kinetoplastida; Trypanosomatidae; Leishmania.
OX NCBI_TaxID=5664;
RN [1]_TaxID=5664;
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=Friedlin;
RA Peacock C.S., Murphy L., Ivens A.C., Berriman M., Blackwell J.,
RA Smith D., Collins M., Foster N., Harris D., Oliver K., O'Neil S.,
RA Saunders D., Seeger K., Warren T., Rajandream M., and Barrell B.G.;
RL Submitted (JUN-2005) to the EMBL/GenBank/DBJ databases.
DR EMBL; C7005272; CAJ09636.1; -; Genomic_DNA.
KW Hypothetical protein.
SQ SEQUENCE 324 AA; 35283 MW; 3D4C39950E516140 CRC64;

Query Match 51.1%; Score 47; DB 2; Length 324;
Best Local Similarity 50.0%; Pred. No. 36;
Matches 9; Conservative 5; Mismatches 4; Indels 0; Gaps 0;

QY 2 SSATTL5AVKADDFDAQI 19
DB 172 SSSSKLSAAEMDDWDADV 189

RESULT 13
FIXC_ECO57
ID FIXC_ECO57 STANDARD; PRT; 428 AA.
AC Q7AHT0; Q8XA26;
DT 01-FEB-2005 (Rel. 46, Created)
DT 01-FEB-2005 (Rel. 46, Last sequence update)
DT 13-SEP-2005 (Rel. 48, Last annotation update)
DE FixC protein.
GN Name=fixC; OrderedLocustNames=z0049, ECs0046;
OS Escherichia coli O157:H7.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Escherichia.
OX NCBI_TaxID=83334;
RN [1]
RP NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
RC STRAIN=O157:H7 / EDL933 / ATCC 700927 / EHEC;
RX MEDLINE=21074935; PubMed=11206551; DOI=10.1038/35054089;
RA Perna N.T., Plunkett G., III, Burland V., Mau B., Glasner J.D.,
RA Rose D.J., Mayhew G.F., Evans P.S., Gregor J., Kirkpatrick H.A.,
RA Posfai G., Hackett J., Klink S., Boutin A., Shao Y., Miller L.,
RA Grobeck E.J., Davis N.W., Lim A., Dimalanta E.T., Potamoudis K.,
RA Apodaca J., Anantharaman T.S., Lin J., Yen G., Schwartz D.C.,
RA Welch R.A., Blattner F.R.;
RT "Genome sequence of enterohaemorrhagic *Escherichia coli* O157:H7.";
RL Nature 409:529-533(2001).
RN [2]
RP NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
RC STRAIN=O157:H7 / Sakai / RIMD 0509952 / EHEC;
RX MEDLINE=21156231; PubMed=11258796;
RA Hayashi T., Makino K., Ohnishi M., Kurokawa K., Ishii K., Yokoyama K.,
RA Han C.-G., Ohtsubo E., Nakayama K., Murata T., Tanaka M., Tobe T.,
RA Iida T., Takami H., Honda T., Sasakawa C., Ogasawara N., Yasunaga T.,
RA Kuhara S., Shiba T., Hattori M., Shinagawa H.;
RT "Complete genome sequence of enterohemorrhagic *Escherichia coli* O157:H7 and genomic comparison with a laboratory strain K-12.";
RL DNA Res. 8:11-22(2001).
CC -!- FUNCTION: Could be part of an electron transfer system required
CC for anaerobic carnitine reduction (By similarity).
CC -!- COFACTOR: FAD (Potential).
CC -!- SIMILARITY: Belongs to the ETF-QO/fixC family.
CC
CC This Swiss-Prot entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use as long as its content is in no way modified and this statement is not
CC removed.
CC
CC EMBL; AE005174; BAG54346.1; -; Genomic_DNA.
DR EMBL; EA000007; BAB33469.1; -; Genomic_DNA.
DR PIR; F85485; F85485.
DR PIR; F90634; F90634.
DR InterPro; IPR006076; FAD_oxred.
DR InterPro; IPR001327; FAD_Pyr_redox.
DR InterPro; IPR000205; NAD_BS.
DR InterPro; IPR001100; Pyr_redox.
DR InterPro; IPR003042; Rng_hydrolyase.
DR Pfam; PF01266; DAO; 1.
DR PRINTS; PR00368; FADPNR.
DR PRINTS; PR00411; PNDRDTASEI.
DR PRINTS; PR00420; RNCMNXGNASE.
KW Complete proteome; Electron transport; FAD; Flavoprotein;
KW Oxidoreductase; Transport.
FT NP_BIND 12 28 FAD or NAD or NADP (ADP part)
FT (Potential).
SQ SEQUENCE 428 AA; 45716 MW; 64B034F30D61DCF1 CRC64;

Query Match 51.1%; Score 47; DB 1; Length 428;
Best Local Similarity 52.9%; Pred. No. 48;

```

Matches 9; Conservative 4; Mismatches 4; Indels 0; Gaps 0;

QY 2 SSATLTSVAKDDPDAQ 18
   ::|||:|||||
Db 327 AAKTVLSAMKSDDFSQ 343

RESULT 14
FIXC_ECOL6
ID FIXC_ECOL6 STANDARD; PRT; 428 AA.
AC P6845; P31575; P75626;
DT 01-JUL-1993 (Rel. 26, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 13-SEP-2005 (Rel. 48, Last annotation update)
DE FixC protein.
GN Name=fixC; OrderedLocusNames=c0053;
OS Escherichia coli O6.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Escherichia.
OX NCBI_TaxID=217992;
RN NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
RC STRAIN=O6:H1 / CFT073 / ATCC 700928 / UPEC;
RX MEDLINE=22388234; PubMed=12471157; DOI=10.1073/pnas.252529799;
RA Welch R.A., Burland V., Plunkett G. III, Redford P., Roesch P.,
RA Rasko D., Buckles E.L., Liou S.-R., Boutin A., Hackett J., Stroud D.,
RA Mayhew G.F., Rose D.J., Zhou S., Schwartz F.R.,
RA Mobley H.L.T., Donnenberg M.S., Blattner F.R.;
RT "Extensive mosaic structure revealed by the complete genome sequence
of uropathogenic Escherichia coli.";
RL Proc. Natl. Acad. Sci. U.S.A. 99:17020-17024(2002).
CC -!- FUNCTION: Could be part of an electron transfer system required
for anaerobic carnitine reduction (By similarity).
CC -!- COFACTOR: FAD (Potential).
CC -!- SIMILARITY: Belongs to the ETP-QO/fixC family.
CC
CC This Swiss-Prot entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use as long as its content is in no way modified and this statement is not
removed.
CC -----
CC EMBL; AB016755; AAN78549.1; -; Genomic_DNA.
CC InterPro; IPR006076; FAD_oxred.
CC InterPro; IPR001327; FAD pyr redox.
CC InterPro; IPR000205; NAD_BS.
CC InterPro; IPR001100; Pyr_redox.
CC InterPro; IPR003042; Rng_hydrolase.
CC Pfam; PF01266; DAO; 1.
CC PRINTS; PR00368; FADPNR.
CC PRINTS; PR00411; PNDRDTASEI.
CC PRINTS; PR00420; RNMNOXGNASE.
KW Complete proteome; Electron transport; FAD; Flavoprotein;
KW Oxidoreductase; Transport.
FT NP_BIND 12 28 FAD or NAD or NADP (ADP part)
SQ SEQUENCE 428 AA; 45703 MW; 4D089EE9A30C7FBC CRC64;

Query Match 51.1%; Score 47; DB 1; Length 428;
Best Local Similarity 52.9%; Pred. No. 48;
Matches 9; Conservative 4; Mismatches 4; Indels 0; Gaps 0;

QY 2 SSATLTSVAKDDPDAQ 18
   ::|||:|||||
Db 327 AAKTVLSAMKSDDFSQ 343

RESULT 15
FIXC_ECOLI
ID FIXC_ECOLI STANDARD; PRT; 428 AA.
AC P68644; P31575; P75626;
DT 01-JUL-1993 (Rel. 26, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)

```

```

DT 13-SEP-2005 (Rel. 48, Last annotation update)
DE FixC protein.
GN Name=fixC; OrderedLocusNames=b0043;
OS Escherichia coli.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Escherichia.
OX NCBI_TaxID=562;
RN NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
RC STRAIN=K12;
RX MEDLINE=92334977; PubMed=1630901;
RA Yura T., Mori H., Nagai H., Nagata T., Ishihama A., Fujita N.,
RA Isono K., Mizobuchi K., Nakata A.;
RT "Systematic sequencing of the Escherichia coli genome: analysis of the
0-2.4 min region.";
RL Nucleic Acids Res. 20:3305-3308(1992).
CC [2]
CC NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
RC STRAIN=K12 / MG1655;
RX MEDLINE=97426617; PubMed=9278503; DOI=10.1126/science.277.5331.1453;
RA Blattner F.R., Plunkett G. III, Bloch C.A., Perna N.T., Burland V.,
RA Riley M., Collado-Vides J., Glasner J.D., Rode C.K., Mayhew G.F.,
RA Gregor J., Davis N.W., Kirkpatrick H.A., Goeden M.A., Rose D.J.,
RA Mau B., Shao Y.;
RT "The complete genome sequence of Escherichia coli K-12.";
RL Science 277:1453-1474(1997).
CC [3]
CC NUCLEOTIDE SEQUENCE OF 1-188.
RC STRAIN=O44:K74;
RX MEDLINE=9606354; PubMed=7473063;
RA Eichler K., Buchet A., Bourgis F., Kleber H.-P.,
RA Mandrand-Berthelot M.-A.;
RT "The fix Escherichia coli region contains four genes related to
carnitine metabolism.";
RL J. Basic Microbiol. 35:217-227(1995).
CC [4]
CC PROBABLE FUNCTION.
RC STRAIN=K12 / BW25113;
RX MEDLINE=22077276; PubMed=12081978;
DOI=10.1128/JB.184.14.4044-4047.2002;
RA Walt A., Kahn M.L.;
RT "The fixA and fixB genes are necessary for anaerobic carnitine
reduction in Escherichia coli.";
RL J. Bacteriol. 184:4044-4047(2002).
CC -!- FUNCTION: Could be part of an electron transfer system required
for anaerobic carnitine reduction.
CC -!- COFACTOR: FAD (Potential).
CC -!- SIMILARITY: Belongs to the ETP-QO/fixC family.
CC -----
CC This Swiss-Prot entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use as long as its content is in no way modified and this statement is not
removed.
CC -----
CC EMBL; D10483; BA96611.1; -; Genomic_DNA.
CC EMBL; U00096; AAC73154.1; -; Genomic_DNA.
CC EMBL; X71977; CAA50799.1; -; Genomic_DNA.
CC PIR; C64725; C64725.
CC EchoBASE; EB1525; -.
CC EcoGene; EG11564; fixC.
CC InterPro; IPR006076; FAD_oxred.
CC InterPro; IPR001327; FAD pyr_redox.
CC InterPro; IPR000205; NAD_BS.
CC InterPro; IPR001100; Pyr_redox.
CC InterPro; IPR003042; Rng_hydrolase.
CC Pfam; PF01266; DAO; 1.
CC PRINTS; PR00368; FADPNR.
CC PRINTS; PR00411; PNDRDTASEI.
CC PRINTS; PR00420; RNMNOXGNASE.
KW Complete proteome; Electron transport; FAD; Flavoprotein;
KW Oxidoreductase; Transport.
FT NP_BIND 12 28 FAD or NAD or NADP (ADP part)

```

```

FT CONFLICT 1 3 (Potential).
FT CONFLICT 64 MSE -> MT (in Ref. 3).
FT CONFLICT 280 64 D -> E (in Ref. 1).
FT CONFLICT 281 281 VV -> IM (in Ref. 1).
FT CONFLICT 310 310 L -> F (in Ref. 1).
SQ SEQUENCE 428 AA; 45703 MW; 4D089EE9A30C7FBC CRC64;

Query Match 51.1%; Score 47; DB 1; Length 428;
Best Local Similarity 52.9%; Pred. No. 48;
Matches 9; Conservative 4; Mismatches 4; Indels 0; Gaps 0;

QY 2 SSATTL$AVKADDFDAQ 18
Db 327 AAKTVLSAMKSDDFSKQ 343
:::|::|::|::|

```

Search completed: January 25, 2006, 19:13:38
 Job time : 71.3 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.1.6
Copyright (c) 1993 - 2006 Compugen Ltd.

OM protein - protein search, using sw model

Run on: January 25, 2006, 18:50:05 ; Search time 16.6 Seconds
(without alignments)
99.609 Million cell updates/sec

Title: US-10-797-821-5
Perfect score: 92
Sequence: 1 LSSATLSAVKADDFDAQIA 20

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 572060 seqs, 82675679 residues

Total number of hits satisfying chosen parameters: 572060

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued Patents AA:*
1: /cgn2_6/prodata/1/iaa/5 COMB.pep.*
2: /cgn2_6/prodata/1/iaa/6 COMB.pep.*
3: /cgn2_6/prodata/1/iaa/H COMB.pep.*
4: /cgn2_6/prodata/1/iaa/ECTUS COMB.pep.*
5: /cgn2_6/prodata/1/iaa/RE COMB.pep.*
6: /cgn2_6/prodata/1/iaa/backfiles1.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	44	47.8	107	2	US-09-270-767-41866
2	44	47.8	1260	2	US-09-328-352-6746
3	43	46.7	525	2	US-09-107-532A-5095
4	43	46.7	1224	2	US-09-901-572A-4
5	42	45.7	675	2	US-09-902-540-14436
6	41	44.6	948	2	US-09-328-352-6123
7	40	43.5	238	2	US-09-602-787A-92
8	40	43.5	505	2	US-09-252-991A-23343
9	40	43.5	556	2	US-09-248-796A-19179
10	40	43.5	614	2	US-10-023-649A-3
11	40	43.5	2492	1	US-09-991-258-3
12	40	43.5	2500	1	US-08-801-263A-2
13	40	43.5	2500	2	US-09-102-248-2
14	40	43.5	2500	2	US-09-367-764-2
15	40	43.5	2512	1	US-08-801-263A-9
16	40	43.5	2512	2	US-09-102-248-9
17	40	43.5	2512	2	US-09-367-764-9
18	40	43.5	2517	1	US-08-801-263A-5
19	40	43.5	2517	2	US-09-102-248-5
20	40	43.5	2517	2	US-09-367-764-5
21	39.5	42.9	223	2	US-10-104-047-2793
22	39	42.4	86	2	US-09-583-110-3661
23	39	42.4	98	2	US-09-107-433-4545
24	39	42.4	155	2	US-09-270-767-32585
25	39	42.4	185	2	US-09-248-796A-21513
26	39	42.4	205	2	US-09-248-796A-19326
27	39	42.4	206	2	US-09-902-540-12815

28 39 42.4 277 2 US-09-602-777A-8
29 39 42.4 327 2 US-09-792-024-83
30 39 42.4 559 2 US-09-328-352-5924
31 39 42.4 621 2 US-08-604-789B-4
32 39 42.4 621 2 US-09-312-721A-4
33 39 42.4 621 2 US-09-733-300-4
34 39 42.4 643 2 US-10-175-002-2
35 39 42.4 755 2 US-09-342-648-2
36 39 42.4 922 2 US-09-198-452A-15
37 39 42.4 932 2 US-09-438-185A-6
38 39 42.4 2431 1 US-07-920-281C-2
39 39 42.4 2431 2 US-08-466-277-2
40 39 42.4 2431 2 US-09-688-842-2
41 38 41.3 150 2 US-09-198-452A-919
42 38 41.3 173 2 US-09-975-132A-28
43 38 41.3 209 2 US-09-543-681A-6765
44 38 41.3 210 2 US-09-583-110-5219
45 38 41.3 218 2 US-09-107-433-4896

ALIGNMENTS

RESULT 1
US-09-270-767-41866
; Sequence 41866, Application US/09270767
; Patent No. 6703491
; GENERAL INFORMATION:
; APPLICANT: Homburger et al.
; TITLE OF INVENTION: Nucleic acids and proteins of Drosophila melanogaster
; FILE REFERENCE: File Reference: 7326-094
; CURRENT APPLICATION NUMBER: US/09/270,767
; CURRENT FILING DATE: 1999-03-17
; NUMBER OF SEQ ID NOS: 62517
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 41866
; LENGTH: 107
; TYPE: PRT
; ORGANISM: Drosophila melanogaster
; FEATURE:
; OTHER INFORMATION: Xaa means any amino acid
US-09-270-767-41866

Query Match 47.8%; Score 44; DB 2; Length 107;
Best Local Similarity 53.3%; Pred. No. 5.8;
Matches 8; Conservative 4; Mismatches 3; Indels 0; Gaps 0;

QY 1 LSSATLSAVKADDF 15
Db 81 IKSATLSLGIRSDHF 95

RESULT 2
US-09-328-352-6746
; Sequence 6746, Application US/09328352
; Patent No. 6562958
; GENERAL INFORMATION:
; APPLICANT: Gary L. Breton et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO ACINETOBACTER
; FILE REFERENCE: GTC99-03PA
; CURRENT APPLICATION NUMBER: US/09/328,352
; CURRENT FILING DATE: 1999-06-04
; NUMBER OF SEQ ID NOS: 8252
; SEQ ID NO 6746
; LENGTH: 1260
; TYPE: PRT
; ORGANISM: Acinetobacter baumannii
US-09-328-352-6746

Query Match 47.8%; Score 44; DB 2; Length 1260;
Best Local Similarity 47.1%; Pred. No. 1.2e+02;
Matches 8; Conservative 3; Mismatches 6; Indels 0; Gaps 0;

Matches 9; Conservative 2; Mismatches 5; Indels 0; Gaps 0;

QY 1 LSSATLSAVKADDFD 16
|:|||||:
Db 668 LMGATTLTAKEKPSID 683

RESULT 7

US-09-602-787A-92
; Sequence 92, Application US/09602787A
; Patent No. 6696561
; GENERAL INFORMATION:
; APPLICANT: Pompejus, Mark
; APPLICANT: Kr"ger, Burkhard
; APPLICANT: Sch"der, Hartwig
; APPLICANT: Zelder, Oskar
; APPLICANT: Habeshauer, Gregor
; TITLE OF INVENTION: CORYNEBACTERIUM GLUTAMICUM GENES ENCODING PROTEINS
; TITLE OF INVENTION: INVOLVED IN MEMBRANE SYNTHESIS AND MEMBRANE
; TITLE OF INVENTION: TRANSPORT
; FILE REFERENCE: BGI-125CP
; CURRENT APPLICATION NUMBER: US/09/602,787A
; CURRENT FILING DATE: 2000-06-23
; PRIOR APPLICATION NUMBER: USSN 60/141031
; PRIOR FILING DATE: 1999-06-25
; PRIOR APPLICATION NUMBER: DE 19931454.3
; PRIOR FILING DATE: 1999-07-08
; PRIOR APPLICATION NUMBER: DE 19931478.0
; PRIOR FILING DATE: 1999-07-08
; PRIOR APPLICATION NUMBER: DE 19931563.9
; PRIOR FILING DATE: 1999-07-08
; PRIOR APPLICATION NUMBER: DE 19932122.1
; PRIOR FILING DATE: 1999-07-09
; PRIOR APPLICATION NUMBER: DE 19932124.8
; PRIOR FILING DATE: 1999-07-09
; PRIOR APPLICATION NUMBER: DE 19932125.6
; PRIOR FILING DATE: 1999-07-09
; PRIOR APPLICATION NUMBER: DE 19932128.0
; PRIOR FILING DATE: 1999-07-09
; PRIOR APPLICATION NUMBER: DE 19932180.9
; PRIOR FILING DATE: 1999-07-09
; PRIOR APPLICATION NUMBER: DE 19932182.5
; PRIOR FILING DATE: 1999-07-09
; PRIOR APPLICATION NUMBER: DE 19932190.6
; PRIOR FILING DATE: 1999-07-09
; PRIOR APPLICATION NUMBER: DE 19932191.4
; PRIOR FILING DATE: 1999-07-09
; PRIOR APPLICATION NUMBER: DE 19932209.0
; PRIOR FILING DATE: 1999-07-09
; PRIOR APPLICATION NUMBER: DE 19932212.0
; PRIOR FILING DATE: 1999-07-09
; PRIOR APPLICATION NUMBER: DE 19932227.9
; PRIOR FILING DATE: 1999-07-09
; PRIOR APPLICATION NUMBER: DE 19932228.7
; PRIOR FILING DATE: 1999-07-09
; PRIOR APPLICATION NUMBER: DE 19932229.5
; PRIOR FILING DATE: 1999-07-09
; PRIOR APPLICATION NUMBER: DE 19932230.9
; PRIOR FILING DATE: 1999-07-09
; PRIOR APPLICATION NUMBER: DE 19932927.3
; PRIOR FILING DATE: 1999-07-14
; PRIOR APPLICATION NUMBER: DE 19933005.0
; PRIOR FILING DATE: 1999-07-14
; PRIOR APPLICATION NUMBER: DE 19933006.9
; PRIOR FILING DATE: 1999-07-14
; PRIOR APPLICATION NUMBER: DE 19940764.9
; PRIOR FILING DATE: 1999-08-27
; PRIOR APPLICATION NUMBER: DE 19940765.7
; PRIOR FILING DATE: 1999-08-27
; PRIOR APPLICATION NUMBER: DE 19940766.5
; PRIOR FILING DATE: 1999-08-27
; PRIOR APPLICATION NUMBER: DE 19940830.0
; PRIOR FILING DATE: 1999-08-27

; PRIOR APPLICATION NUMBER: DE 19940831.9
; PRIOR FILING DATE: 1999-08-27
; PRIOR APPLICATION NUMBER: DE 19940832.7
; PRIOR FILING DATE: 1999-08-27
; PRIOR APPLICATION NUMBER: DE 19940833.5
; PRIOR FILING DATE: 1999-08-27
; PRIOR APPLICATION NUMBER: DE 19941378.9
; PRIOR FILING DATE: 1999-08-31
; PRIOR APPLICATION NUMBER: DE 19941379.7
; PRIOR FILING DATE: 1999-08-31
; PRIOR APPLICATION NUMBER: DE 19941395.9
; PRIOR FILING DATE: 1999-08-31
; PRIOR APPLICATION NUMBER: DE 19942077.7
; PRIOR FILING DATE: 1999-09-03
; PRIOR APPLICATION NUMBER: DE 19942078.5
; PRIOR FILING DATE: 1999-09-03
; PRIOR APPLICATION NUMBER: DE 19942079.3
; PRIOR FILING DATE: 1999-09-03
; PRIOR APPLICATION NUMBER: DE 19942088.2
; PRIOR FILING DATE: 1999-09-03
; NUMBER OF SEQ ID NOS: 678
; SEQ ID NO 92
; LENGTH: 238
; TYPE: PRT
; ORGANISM: Corynebacterium glutamicum
US-09-602-787A-92

Query Match 43.5%; Score 40; DB 2; Length 238;
Best Local Similarity 55.6%; Pred. No. 71;
Matches 10; Conservative 2; Mismatches 6; Indels 0; Gaps 0;

QY 3 SATLSAVKADDFDAQIA 20
:|||||:
Db 77 AAQTLSAETADDEDTVLA 94

RESULT 8

US-09-252-991A-29343
; Sequence 29343, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 29343
; LENGTH: 505
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-29343

Query Match 43.5%; Score 40; DB 2; Length 505;
Best Local Similarity 50.0%; Pred. No. 1.8e+02;
Matches 7; Conservative 4; Mismatches 3; Indels 0; Gaps 0;

QY 7 LSAVKADDFDAQIA 20
|:::|:
Db 147 LASVRLDLDLQVA 160

RESULT 9
US-09-248-796A-19179
; Sequence 19179, Application US/09248796A
; Patent No. 6747137
; GENERAL INFORMATION:
; APPLICANT: Keith Weinstock et al

```

; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO CANDIDA ALBICAN
; TITLE OF INVENTION: FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 107196.132
; CURRENT APPLICATION NUMBER: US/09/248,796A
; CURRENT FILING DATE: 1999-02-12
; PRIOR APPLICATION NUMBER: US 60/074,725
; PRIOR FILING DATE: 1998-02-13
; PRIOR APPLICATION NUMBER: US 60/096,409
; PRIOR FILING DATE: 1998-08-13
; NUMBER OF SEQ ID NOS: 28208
; SEQ ID NO 19179
; LENGTH: 556
; TYPE: PRT
; ORGANISM: Candida albicans
; US-09-248-796A-19179

Query Match      43.5%; Score 40; DB 2; Length 556;
Best Local Similarity 50.0%; Pred. No. 2e+02;
Matches 7; Conservative 4; Mismatches 3; Indels 0; Gaps 0;

QY      7 LSAVKADDFDAQIA 20
      | : | : | : | : |
Db      104 LEMIQAPDFDSQLA 117

RESULT 10
US-10-023-649A-3
; Sequence 3, Application US/10023649A
; Patent No. 6800289
; GENERAL INFORMATION:
; APPLICANT: Nagata, Leslie P.
; APPLICANT: Wong, Jonathan P.
; TITLE OF INVENTION: A STRAIN OF THE WESTERN EQUINE ENCEPHALITIS VIRUS (AS AMENDED)
; FILE REFERENCE: NEL-001
; CURRENT APPLICATION NUMBER: US/10/023,649A
; CURRENT FILING DATE: 2001-12-21
; PRIOR APPLICATION NUMBER: 60/256,948
; PRIOR FILING DATE: 2000-12-21
; NUMBER OF SEQ ID NOS: 49
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 3
; LENGTH: 614
; TYPE: PRT
; ORGANISM: Western equine encephalomyelitis virus - strain 71V-1659
; US-10-023-649A-3

Query Match      43.5%; Score 40; DB 2; Length 614;
Best Local Similarity 50.0%; Pred. No. 2.3e+02;
Matches 10; Conservative 3; Mismatches 7; Indels 0; Gaps 0;

QY      1 LSSATTL SAVKADDFDAQIA 20
      | : | : | : | : |
Db      345 LPNIHTLFDMSAEDFDAIIA 364

RESULT 11
US-09-991-258-3
; Sequence 3, Application US/09991258
; Patent No. 6783939
; GENERAL INFORMATION:
; APPLICANT: Olmsted, Robert
; APPLICANT: Keith, Paula
; APPLICANT: Dryga, Sergey
; APPLICANT: Caley, Ian
; APPLICANT: Maughan, Maureen
; APPLICANT: Johnston, Robert
; APPLICANT: Davis, Nancy
; APPLICANT: Swanstrom, Ronald
; TITLE OF INVENTION: ALPHAVIRUS VECTORS AND VIROSOMES WITH MODIFIED HIV GENES FOR USE
; TITLE OF INVENTION: VACCINES
; FILE REFERENCE: 01113.0001U3
; CURRENT APPLICATION NUMBER: US/09/991,258
; CURRENT FILING DATE: 2001-11-16

```

```

; PRIOR APPLICATION NUMBER: 09/902,537
; PRIOR FILING DATE: 2001-07-09
; PRIOR APPLICATION NUMBER: 60/216,995
; PRIOR FILING DATE: 2000-07-07
; NUMBER OF SEQ ID NOS: 19
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 3
; LENGTH: 2492
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence; No. 6783939e =
; OTHER INFORMATION: synthetic construct
; US-09-991-258-3

Query Match      43.5%; Score 40; DB 2; Length 2492;
Best Local Similarity 50.0%; Pred. No. 1.2e+03;
Matches 10; Conservative 3; Mismatches 7; Indels 0; Gaps 0;

QY      1 LSSATTL SAVKADDFDAQIA 20
      | : | : | : | : |
Db      223 LPNIHTLFDMSAEDFDAIIA 2242

RESULT 12
US-08-801-263A-2
; Sequence 2, Application US/08801263A
; Patent No. 5811407
; GENERAL INFORMATION:
; APPLICANT: Johnston, Robert E.
; APPLICANT: Davis, Nancy L.
; APPLICANT: Simpson, Dennis A.
; TITLE OF INVENTION: System for the In Vivo Delivery and
; TITLE OF INVENTION: Expression of Heterologous Genes in the Bone Marrow
; NUMBER OF SEQUENCES: 12
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Bell Seltzer Park & Gibson, P.A.
; STREET: 1211 East Morehead Street
; CITY: Charlotte
; STATE: No. 5811407th Carolina
; COUNTRY: USA
; ZIP: 28234
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/801,263A
; FILING DATE: 19-FEB-1997
; CLASSIFICATION: 514
; ATTORNEY/AGENT INFORMATION:
; NAME: Sibley, Kenneth D.
; REGISTRATION NUMBER: 31,665
; REFERENCE/DOCKET NUMBER: 5470-147
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 919-420-2200
; TELEFAX: 919-881-3175
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2500 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-801-263A-2

Query Match      43.5%; Score 40; DB 1; Length 2500;
Best Local Similarity 50.0%; Pred. No. 1.3e+03;
Matches 10; Conservative 3; Mismatches 7; Indels 0; Gaps 0;

QY      1 LSSATTL SAVKADDFDAQIA 20
      | : | : | : | : |
Db      2228 LPNIHTLFDMSAEDFDAIIA 2247

```

```
RESULT 13
US-09-102-248-2
; Sequence 2, Application US/09102248
; Patent No. 6008035
; GENERAL INFORMATION:
; APPLICANT: Johnston, Robert E.
; APPLICANT: Davis, Nancy L.
; APPLICANT: Simpson, Dennis A.
; TITLE OF INVENTION: System for the In Vivo Delivery and
; NUMBER OF SEQUENCES: 12
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Bell Seltzer Park & Gibson, P.A.
; STREET: 1211 East Morehead Street
; CITY: Charlotte
; STATE: No. 6008035th Carolina
; COUNTRY: USA
; ZIP: 28234
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/102,248
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/801,263
; FILING DATE: 19-FEB-1997
; ATTORNEY/AGENT INFORMATION:
; NAME: Sibley, Kenneth D.
; REGISTRATION NUMBER: 31,665
; REFERENCE/DOCKET NUMBER: 5470-147
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 919-420-2200
; TELEFAX: 919-881-3175
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2500 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-09-102-248-2

Query Match 43.5%; Score 40; DB 2; Length 2500;
Best Local Similarity 50.0%; Pred. No. 1.3e+03;
Matches 10; Conservative 3; Mismatches 7; Indels 0; Gaps 0;

QY 1 LSSATLTSVAKADDFDAQIA 20
| : | : | : | : | : |
Db 2228 LPNIHTLPDMSAEDFDAIIA 2247

RESULT 14
US-09-367-764-2
; Sequence 2, Application US/09367764
; Patent No. 6583121
; GENERAL INFORMATION:
; APPLICANT: Johnston, Robert E.
; APPLICANT: Davis, Nancy L.
; APPLICANT: Simpson, Dennis A.
; TITLE OF INVENTION: System for the In Vivo Delivery and
; NUMBER OF SEQUENCES: 12
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Bell Seltzer Park & Gibson, P.A.
; STREET: 1211 East Morehead Street
; CITY: Charlotte
; STATE: No. 6583121th Carolina
; COUNTRY: USA

Query Match 43.5%; Score 40; DB 2; Length 2500;
Best Local Similarity 50.0%; Pred. No. 1.3e+03;
Matches 10; Conservative 3; Mismatches 7; Indels 0; Gaps 0;

QY 1 LSSATLTSVAKADDFDAQIA 20
| : | : | : | : | : |
Db 2228 LPNIHTLPDMSAEDFDAIIA 2247

RESULT 15
US-08-801-263A-9
; Sequence 9, Application US/08801263A
; Patent No. 5811407
; GENERAL INFORMATION:
; APPLICANT: Johnston, Robert E.
; APPLICANT: Davis, Nancy L.
; APPLICANT: Simpson, Dennis A.
; TITLE OF INVENTION: System for the In Vivo Delivery and
; NUMBER OF SEQUENCES: 12
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Bell Seltzer Park & Gibson, P.A.
; STREET: 1211 East Morehead Street
; CITY: Charlotte
; STATE: No. 5811407th Carolina
; COUNTRY: USA
; ZIP: 28234
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/801,263A
; FILING DATE: 19-FEB-1997
; CLASSIFICATION: 514
; ATTORNEY/AGENT INFORMATION:
; NAME: Sibley, Kenneth D.
; REGISTRATION NUMBER: 31,665
; REFERENCE/DOCKET NUMBER: 5470-147
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 919-420-2200
; TELEFAX: 919-881-3175
; INFORMATION FOR SEQ ID NO: 9:
; SEQUENCE CHARACTERISTICS:
```

```
; LENGTH: 2512 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-801-263A-9
```

```
Query Match      43.5%; Score 40; DB 1; Length 2512;
Best Local Similarity 50.0%; Pred. No. 1.3e+03;
Matches 10; Conservative 3; Mismatches 7; Indels 0; Gaps 0;
```

```
QY      1 LSSATTL SAVKADDFDAQIA 20
      | : | | : | : | | | |
Db      2240 LPNIHTLPDMSAEDFDALIA 2259
```

```
Search completed: January 25, 2006, 19:18:13
Job time : 16.6 secs
```

GenCore version 5.1.6
Copyright (c) 1993 - 2006 Compugen Ltd.

OM protein - protein search, using sw model

Run on: January 25, 2006, 19:13:51 ; Search time 60 Seconds
(without alignments)
139.276 Million cell updates/sec

Title: US-10-797-821-5
Perfect score: 92
Sequence: 1 LSSATTL SAVKADDFDAQIA 20

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1867569 seqs, 417829326 residues

Total number of hits satisfying chosen parameters: 1867569

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published Applications AA Main: *
1: /cgn2_6/ptodata/1/pubpaa/US07_PUBCOMB.pep.*
2: /cgn2_6/ptodata/1/pubpaa/US08_PUBCOMB.pep.*
3: /cgn2_6/ptodata/1/pubpaa/US09_PUBCOMB.pep.*
4: /cgn2_6/ptodata/1/pubpaa/US10A_PUBCOMB.pep.*
5: /cgn2_6/ptodata/1/pubpaa/US10B_PUBCOMB.pep.*
6: /cgn2_6/ptodata/1/pubpaa/US11_PUBCOMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	92	100.0	20	4	US-10-383-930-5
2	92	100.0	20	5	US-10-797-821-5
3	92	100.0	431	4	US-10-383-930-29
4	92	100.0	431	4	US-10-383-930-30
5	92	100.0	431	4	US-10-383-930-33
6	92	100.0	431	5	US-10-797-821-29
7	92	100.0	431	5	US-10-797-821-30
8	92	100.0	431	5	US-10-797-821-33
9	92	100.0	432	4	US-10-383-930-32
10	92	100.0	432	5	US-10-797-821-32
11	91	98.9	432	4	US-10-383-930-31
12	91	98.9	432	5	US-10-797-821-31
13	46	50.0	694	6	US-11-097-143-38595
14	46	50.0	443	6	US-11-097-143-41619
15	44	47.8	446	4	US-10-282-122A-51310
16	43	46.7	380	4	US-10-369-493-3145
17	43	46.7	483	4	US-10-425-114-39646
18	43	46.7	524	4	US-10-282-122A-57658
19	43	46.7	1042	4	US-10-131-591A-78
20	43	46.7	1224	3	US-09-901-572A-4
21	43	46.7	1224	4	US-10-131-591A-25
22	43	46.7	1432	4	US-10-282-122A-52639
23	42	45.7	20	4	US-10-383-930-4
24	42	45.7	20	5	US-10-797-821-4
25	42	45.7	67	4	US-10-437-963-113160
26	42	45.7	113	4	US-10-424-599-252326
27	42	45.7	113	4	US-10-437-963-158213

28	42	45.7	227	4	US-10-425-115-250586	Sequence 250586,
29	42	45.7	427	4	US-10-369-493-769	Sequence 769, App
30	42	45.7	427	4	US-10-282-122A-43186	Sequence 43186, A
31	41	44.6	127	6	US-11-097-143-9270	Sequence 9270, Ap
32	41	44.6	310	6	US-11-097-143-14547	Sequence 14547, A
33	41	44.6	705	4	US-10-424-599-194501	Sequence 194501,
34	41	44.6	878	4	US-10-437-963-164591	Sequence 164591,
35	41	44.6	943	4	US-10-282-122A-44848	Sequence 44848, A
36	40	43.5	127	4	US-10-425-115-266224	Sequence 266224,
37	40	43.5	153	4	US-10-767-701-45045	Sequence 45045, A
38	40	43.5	181	4	US-10-437-963-121684	Sequence 121684,
39	40	43.5	186	4	US-10-369-493-21797	Sequence 21797, A
40	40	43.5	190	4	US-10-108-260A-3234	Sequence 3234, Ap
41	40	43.5	223	3	US-09-738-626-3860	Sequence 3860, Ap
42	40	43.5	238	4	US-10-627-476-92	Sequence 92, Appl
43	40	43.5	254	4	US-10-369-493-12323	Sequence 12323, A
44	40	43.5	271	4	US-10-156-761-8135	Sequence 8135, Ap
45	40	43.5	323	4	US-10-767-701-39937	Sequence 39937, A

ALIGNMENTS

RESULT 1
US-10-383-930-5
; Sequence 5, Application US/10383930
; Publication No. US20040127400A1
; GENERAL INFORMATION:
; APPLICANT: Smith, Daniel J
; APPLICANT: Taubman, Martin A
; TITLE OF INVENTION: Immunogenicity of Glucan Binding Protein
; FILE REFERENCE: 25669-018
; CURRENT APPLICATION NUMBER: US/10/383,930
; CURRENT FILING DATE: 2003-03-07
; PRIOR APPLICATION NUMBER: 60/402,483
; PRIOR FILING DATE: 2002-08-08
; PRIOR APPLICATION NUMBER: 60/363,209
; PRIOR FILING DATE: 2002-03-07
; NUMBER OF SEQ ID NOS: 41
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 5
; LENGTH: 20
; TYPE: PRT
; ORGANISM: Streptococcus mutans
US-10-383-930-5

Query Match 100.0%; Score 92; DB 4; Length 20;
Best Local Similarity 100.0%; Pred. No. 3.7e-08; Mismatches 0; Indels 0; Gaps 0;
Matches 20; Conservative 0;
QY 1 LSSATTL SAVKADDFDAQIA 20
|||||
Db 1 LSSATTL SAVKADDFDAQIA 20

RESULT 2
US-10-797-821-5
; Sequence 5, Application US/10797821
; Publication No. US20050031633A1
; GENERAL INFORMATION:
; APPLICANT: Smith, Daniel J.
; APPLICANT: Taubman, Martin A.
; TITLE OF INVENTION: Glucan Binding Protein and Glycosyltransferase Immunogens
; FILE REFERENCE: 25669-020
; CURRENT APPLICATION NUMBER: US/10/797,821
; CURRENT FILING DATE: 2004-03-09
; PRIOR APPLICATION NUMBER: 10/383,930
; PRIOR FILING DATE: 2003-03-07
; PRIOR APPLICATION NUMBER: 60/363,209
; PRIOR FILING DATE: 2002-03-07
; PRIOR APPLICATION NUMBER: 60/402,483
; PRIOR FILING DATE: 2002-08-08
; PRIOR APPLICATION NUMBER: 09/290,049

; PRIOR FILING DATE: 1999-04-12
; PRIOR APPLICATION NUMBER: 60/081,550
; PRIOR FILING DATE: 1998-04-13
; PRIOR APPLICATION NUMBER: 60/115,142
; PRIOR FILING DATE: 1999-01-08
; NUMBER OF SEQ ID NOS: 45
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 5
; LENGTH: 20
; TYPE: PRT
; ORGANISM: Artificial
; FEATURE:
; OTHER INFORMATION: GbpB peptide
US-10-797-821-5

Query Match 100.0%; Score 92; DB 5; Length 20;
Best Local Similarity 100.0%; Pred. No. 3.7e-08;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 LSSATTLSAVKADDFDAQIA 20
|||||
Db 1 LSSATTLSAVKADDFDAQIA 20
|||||

RESULT 3
US-10-383-930-29
; Sequence 29, Application US/10383930
; Publication No. US20040127400A1
; GENERAL INFORMATION:
; APPLICANT: Smith, Daniel J
; TITLE OF INVENTION: Immunogenicity of Glucan Binding Protein
; FILE REFERENCE: 25669-018
; CURRENT APPLICATION NUMBER: US/10/383,930
; CURRENT FILING DATE: 2003-03-07
; PRIOR APPLICATION NUMBER: 60/402,483
; PRIOR FILING DATE: 2002-08-08
; PRIOR APPLICATION NUMBER: 60/363,209
; PRIOR FILING DATE: 2002-03-07
; NUMBER OF SEQ ID NOS: 41
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 29
; LENGTH: 431
; TYPE: PRT
; ORGANISM: Streptococcus mutans
US-10-383-930-29

Query Match 100.0%; Score 92; DB 4; Length 431;
Best Local Similarity 100.0%; Pred. No. 1.4e-06;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 LSSATTLSAVKADDFDAQIA 20
|||||
Db 16 LSSATTLSAVKADDFDAQIA 35
|||||

RESULT 4
US-10-383-930-30
; Sequence 30, Application US/10383930
; Publication No. US20040127400A1
; GENERAL INFORMATION:
; APPLICANT: Smith, Daniel J
; TITLE OF INVENTION: Immunogenicity of Glucan Binding Protein
; FILE REFERENCE: 25669-018
; CURRENT APPLICATION NUMBER: US/10/383,930
; CURRENT FILING DATE: 2003-03-07
; PRIOR APPLICATION NUMBER: 60/402,483
; PRIOR FILING DATE: 2002-08-08
; PRIOR APPLICATION NUMBER: 60/363,209
; PRIOR FILING DATE: 2002-03-07
; NUMBER OF SEQ ID NOS: 41
; SOFTWARE: PatentIn version 3.2

; SEQ ID NO 30
; LENGTH: 431
; TYPE: PRT
; ORGANISM: Streptococcus mutans
US-10-383-930-30

Query Match 100.0%; Score 92; DB 4; Length 431;
Best Local Similarity 100.0%; Pred. No. 1.4e-06;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 LSSATTLSAVKADDFDAQIA 20
|||||
Db 16 LSSATTLSAVKADDFDAQIA 35
|||||

RESULT 5
US-10-383-930-33
; Sequence 33, Application US/10383930
; Publication No. US20040127400A1
; GENERAL INFORMATION:
; APPLICANT: Smith, Daniel J
; TITLE OF INVENTION: Immunogenicity of Glucan Binding Protein
; FILE REFERENCE: 25669-018
; CURRENT APPLICATION NUMBER: US/10/383,930
; CURRENT FILING DATE: 2003-03-07
; PRIOR APPLICATION NUMBER: 60/402,483
; PRIOR FILING DATE: 2002-08-08
; PRIOR APPLICATION NUMBER: 60/363,209
; PRIOR FILING DATE: 2002-03-07
; NUMBER OF SEQ ID NOS: 41
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 33
; LENGTH: 431
; TYPE: PRT
; ORGANISM: Streptococcus mutans
US-10-383-930-33

Query Match 100.0%; Score 92; DB 4; Length 431;
Best Local Similarity 100.0%; Pred. No. 1.4e-06;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 LSSATTLSAVKADDFDAQIA 20
|||||
Db 16 LSSATTLSAVKADDFDAQIA 35
|||||

RESULT 6
US-10-797-821-29
; Sequence 29, Application US/10797821
; Publication No. US20050031633A1
; GENERAL INFORMATION:
; APPLICANT: Smith, Daniel J.
; TITLE OF INVENTION: Glucan Binding Protein and Glycosyltransferase Immunogens
; FILE REFERENCE: 25669-020
; CURRENT APPLICATION NUMBER: US/10/797,821
; CURRENT FILING DATE: 2004-03-09
; PRIOR APPLICATION NUMBER: 10/383,930
; PRIOR FILING DATE: 2003-03-07
; PRIOR APPLICATION NUMBER: 60/363,209
; PRIOR FILING DATE: 2002-03-07
; PRIOR APPLICATION NUMBER: 60/402,483
; PRIOR FILING DATE: 2002-08-08
; PRIOR APPLICATION NUMBER: 09/290,049
; PRIOR FILING DATE: 1999-04-12
; PRIOR APPLICATION NUMBER: 60/081,550
; PRIOR FILING DATE: 1998-04-13
; PRIOR APPLICATION NUMBER: 60/115,142
; PRIOR FILING DATE: 1999-01-08
; NUMBER OF SEQ ID NOS: 45
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 29

```
; LENGTH: 431
; TYPE: PRT
; ORGANISM: Streptococcus mutans
US-10-797-821-29

Query Match      100.0%; Score 92; DB 5; Length 431;
Best Local Similarity 100.0%; Pred. No. 1.4e-06;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 LSSATTLSAVKADDFDAQIA 20
Db 16 LSSATTLSAVKADDFDAQIA 35

RESULT 7
US-10-797-821-30
; Sequence 30, Application US/10797821
; Publication No. US20050031633A1
; GENERAL INFORMATION:
; APPLICANT: Smith, Daniel J.
; APPLICANT: Taubman, Martin A.
; TITLE OF INVENTION: Glucan Binding Protein and Glycosyltransferase Immunogens
; FILE REFERENCE: 25669-020
; CURRENT FILING DATE: 2004-03-09
; PRIOR APPLICATION NUMBER: US/10/797,821
; PRIOR FILING DATE: 2003-03-07
; PRIOR APPLICATION NUMBER: 10/383,930
; PRIOR FILING DATE: 2002-03-07
; PRIOR APPLICATION NUMBER: 60/363,209
; PRIOR FILING DATE: 2002-03-07
; PRIOR APPLICATION NUMBER: 60/402,483
; PRIOR FILING DATE: 2002-08-08
; PRIOR APPLICATION NUMBER: 09/290,049
; PRIOR FILING DATE: 1999-04-12
; PRIOR APPLICATION NUMBER: 60/081,550
; PRIOR FILING DATE: 1998-04-13
; PRIOR APPLICATION NUMBER: 60/115,142
; PRIOR FILING DATE: 1999-01-08
; NUMBER OF SEQ ID NOS: 45
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 30
; LENGTH: 431
; TYPE: PRT
; ORGANISM: Streptococcus mutans
US-10-797-821-30

Query Match      100.0%; Score 92; DB 5; Length 431;
Best Local Similarity 100.0%; Pred. No. 1.4e-06;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 LSSATTLSAVKADDFDAQIA 20
Db 16 LSSATTLSAVKADDFDAQIA 35

RESULT 8
US-10-797-821-33
; Sequence 33, Application US/10797821
; Publication No. US20050031633A1
; GENERAL INFORMATION:
; APPLICANT: Smith, Daniel J.
; APPLICANT: Taubman, Martin A.
; TITLE OF INVENTION: Glucan Binding Protein and Glycosyltransferase Immunogens
; FILE REFERENCE: 25669-020
; CURRENT FILING DATE: 2004-03-09
; PRIOR APPLICATION NUMBER: US/10/797,821
; PRIOR FILING DATE: 2003-03-07
; PRIOR APPLICATION NUMBER: 10/383,930
; PRIOR FILING DATE: 2002-03-07
; PRIOR APPLICATION NUMBER: 60/363,209
; PRIOR FILING DATE: 2002-03-07
; PRIOR APPLICATION NUMBER: 60/402,483
; PRIOR FILING DATE: 2002-08-08
; PRIOR APPLICATION NUMBER: 09/290,049
; PRIOR FILING DATE: 1999-04-12
; PRIOR APPLICATION NUMBER: 60/081,550
; PRIOR FILING DATE: 1998-04-13
; PRIOR APPLICATION NUMBER: 60/115,142
; PRIOR FILING DATE: 1999-01-08
; NUMBER OF SEQ ID NOS: 45
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 33
; LENGTH: 431
; TYPE: PRT
; ORGANISM: Streptococcus mutans
US-10-797-821-33

Query Match      100.0%; Score 92; DB 5; Length 431;
Best Local Similarity 100.0%; Pred. No. 1.4e-06;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 LSSATTLSAVKADDFDAQIA 20
Db 16 LSSATTLSAVKADDFDAQIA 35

RESULT 9
US-10-383-930-32
; Sequence 32, Application US/10383930
; Publication No. US20040127400A1
; GENERAL INFORMATION:
; APPLICANT: Smith, Daniel J.
; APPLICANT: Taubman, Martin A.
; TITLE OF INVENTION: Immunogenicity of Glucan Binding Protein
; FILE REFERENCE: 25669-018
; CURRENT APPLICATION NUMBER: US/10/383,930
; CURRENT FILING DATE: 2003-03-07
; PRIOR APPLICATION NUMBER: 60/402,483
; PRIOR FILING DATE: 2002-08-08
; PRIOR APPLICATION NUMBER: 60/363,209
; PRIOR FILING DATE: 2002-03-07
; NUMBER OF SEQ ID NOS: 41
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 32
; LENGTH: 432
; TYPE: PRT
; ORGANISM: Streptococcus mutans
US-10-383-930-32

Query Match      100.0%; Score 92; DB 4; Length 432;
Best Local Similarity 100.0%; Pred. No. 1.5e-06;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 LSSATTLSAVKADDFDAQIA 20
Db 16 LSSATTLSAVKADDFDAQIA 35

RESULT 10
US-10-797-821-32
; Sequence 32, Application US/10797821
; Publication No. US20050031633A1
; GENERAL INFORMATION:
; APPLICANT: Smith, Daniel J.
; APPLICANT: Taubman, Martin A.
; TITLE OF INVENTION: Glucan Binding Protein and Glycosyltransferase Immunogens
; FILE REFERENCE: 25669-020
; CURRENT FILING DATE: 2004-03-09
; PRIOR APPLICATION NUMBER: US/10/797,821
; PRIOR FILING DATE: 2003-03-07
; PRIOR APPLICATION NUMBER: 10/383,930
; PRIOR FILING DATE: 2002-03-07
; PRIOR APPLICATION NUMBER: 60/363,209
; PRIOR FILING DATE: 2002-03-07
; PRIOR APPLICATION NUMBER: 60/402,483
; PRIOR FILING DATE: 2002-08-08
; PRIOR APPLICATION NUMBER: 09/290,049
; PRIOR FILING DATE: 1999-04-12
; PRIOR APPLICATION NUMBER: 60/081,550
```

```
; PRIOR FILING DATE: 1998-04-13
; PRIOR APPLICATION NUMBER: 60/115,142
; PRIOR FILING DATE: 1999-01-08
; NUMBER OF SEQ ID NOS: 45
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 32
; LENGTH: 432
; TYPE: PRT
; ORGANISM: Streptococcus mutans
US-10-797-821-32
```

```
Query Match 100.0%; Score 92; DB 5; Length 432;
Best Local Similarity 100.0%; Pred. No. 1.5e-06;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY 1 LSSATTL SAVKADDFDAQIA 20
| | | | | | | | | | | | | | | | | |
DB 16 LSSATTL SAVKADDFDAQIA 35
```

```
RESULT 11
US-10-383-930-31
; Sequence 31, Application US/10383930
; Publication No. US20040127400A1
; GENERAL INFORMATION:
; APPLICANT: Smith, Daniel J
; APPLICANT: Taubman, Martin A
; TITLE OF INVENTION: Immunogenicity of Glucan Binding Protein
; FILE REFERENCE: 25669-018
; CURRENT APPLICATION NUMBER: US/10/383,930
; CURRENT FILING DATE: 2003-03-07
; PRIOR APPLICATION NUMBER: 60/402,483
; PRIOR FILING DATE: 2002-08-08
; PRIOR APPLICATION NUMBER: 60/363,209
; PRIOR FILING DATE: 2002-03-07
; NUMBER OF SEQ ID NOS: 41
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 31
; LENGTH: 432
; TYPE: PRT
; ORGANISM: Streptococcus mutans
US-10-383-930-31
```

```
Query Match 98.9%; Score 91; DB 4; Length 432;
Best Local Similarity 95.0%; Pred. No. 2.1e-06;
Matches 19; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
```

```
QY 1 LSSATTL SAVKADDFDAQIA 20
| | | | | | | | | | | | | | | | | |
DB 16 LSSATTL SAVKADDFDAQIA 35
```

```
RESULT 12
US-10-797-821-31
; Sequence 31, Application US/10797821
; Publication No. US20050031633A1
; GENERAL INFORMATION:
; APPLICANT: Smith, Daniel J.
; APPLICANT: Taubman, Martin A.
; TITLE OF INVENTION: Glucan Binding Protein and Glycosyltransferase Immunogens
; FILE REFERENCE: 25669-020
; CURRENT APPLICATION NUMBER: US/10/797,821
; CURRENT FILING DATE: 2004-03-09
; PRIOR APPLICATION NUMBER: 10/383,930
; PRIOR FILING DATE: 2003-03-07
; PRIOR APPLICATION NUMBER: 60/363,209
; PRIOR FILING DATE: 2002-03-07
; PRIOR APPLICATION NUMBER: 60/402,483
; PRIOR FILING DATE: 2002-08-08
; PRIOR APPLICATION NUMBER: 09/250,049
; PRIOR FILING DATE: 1999-04-12
; PRIOR APPLICATION NUMBER: 60/081,550
; PRIOR FILING DATE: 1998-04-13
```

```
; PRIOR APPLICATION NUMBER: 60/115,142
; PRIOR FILING DATE: 1999-01-08
; NUMBER OF SEQ ID NOS: 45
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 31
; LENGTH: 432
; TYPE: PRT
; ORGANISM: Streptococcus mutans
US-10-797-821-31
```

```
Query Match 98.9%; Score 91; DB 5; Length 432;
Best Local Similarity 95.0%; Pred. No. 2.1e-06;
Matches 19; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
```

```
QY 1 LSSATTL SAVKADDFDAQIA 20
| | | | | | | | | | | | | | | | | |
DB 16 LSSATTL SAVKADDFDAQIA 35
```

```
RESULT 13
US-11-097-143-38595
; Sequence 38595, Application US/11097143
; Publication No. US20050208558A1
; GENERAL INFORMATION:
; APPLICANT: Venter, J. Craig
; APPLICANT: et al.
; TITLE OF INVENTION: DETECTION KIT, SUCH AS NUCLEIC ACID
; TITLE OF INVENTION: ARRAYS, FOR DETECTING EXPRESSION OF 10,000 OR MORE
; FILE REFERENCE: CL000728
; CURRENT APPLICATION NUMBER: US/11/097,143
; CURRENT FILING DATE: 2005-04-04
; PRIOR APPLICATION NUMBER: 60/157,832
; PRIOR FILING DATE: 1999-10-05
; PRIOR APPLICATION NUMBER: 60/160,191
; PRIOR FILING DATE: 1999-10-19
; PRIOR APPLICATION NUMBER: 60/161,932
; PRIOR FILING DATE: 1999-10-28
; PRIOR APPLICATION NUMBER: 60/164,769
; PRIOR FILING DATE: 1999-11-12
; PRIOR APPLICATION NUMBER: 60/173,383
; PRIOR FILING DATE: 1999-12-28
; PRIOR APPLICATION NUMBER: 60/175,693
; PRIOR FILING DATE: 2000-01-12
; PRIOR APPLICATION NUMBER: 60/184,831
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: 60/191,637
; PRIOR FILING DATE: 2000-03-23
; NUMBER OF SEQ ID NOS: 43008
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 38595
; LENGTH: 694
; TYPE: PRT
; ORGANISM: DROSOPHILA
US-11-097-143-38595
```

```
Query Match 50.0%; Score 46; DB 6; Length 694;
Best Local Similarity 64.3%; Pred. No. 1e+02;
Matches 9; Conservative 2; Mismatches 3; Indels 0; Gaps 0;
```

```
QY 5 TTLSAVKADDFDAQ 18
| | | | | | | | | | | | | | | | | |
DB 499 TLPSPRADDFDAQ 512
```

```
RESULT 14
US-11-097-143-41619
; Sequence 41619, Application US/11097143
; Publication No. US20050208558A1
; GENERAL INFORMATION:
; APPLICANT: Venter, J. Craig
; APPLICANT: et al.
; TITLE OF INVENTION: DETECTION KIT, SUCH AS NUCLEIC ACID
```



```

; TITLE OF INVENTION: ARRAYS, FOR DETECTING EXPRESSION OF 10,000 OR MORE
; FILE REFERENCE: CLO000728
; CURRENT APPLICATION NUMBER: US/11/097,143
; CURRENT FILING DATE: 2005-04-04
; PRIOR APPLICATION NUMBER: 60/157,832
; PRIOR FILING DATE: 1999-10-05
; PRIOR APPLICATION NUMBER: 60/160,191
; PRIOR FILING DATE: 1999-10-19
; PRIOR APPLICATION NUMBER: 60/161,932
; PRIOR FILING DATE: 1999-10-28
; PRIOR APPLICATION NUMBER: 60/164,769
; PRIOR FILING DATE: 1999-11-12
; PRIOR APPLICATION NUMBER: 60/173,383
; PRIOR FILING DATE: 1999-12-28
; PRIOR APPLICATION NUMBER: 60/175,693
; PRIOR FILING DATE: 2000-01-12
; PRIOR APPLICATION NUMBER: 60/184,831
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: 60/191,637
; PRIOR FILING DATE: 2000-03-23
; NUMBER OF SEQ ID NOS: 43008
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 41619
; LENGTH: 4643
; TYPE: PRT
; ORGANISM: DROSOPHILA
US-11-097-143-41619

Query Match          50.0%; Score 46; DB 6; Length 4643;
Best Local Similarity 61.1%; Pred. No. 1e+03;
Matches 11; Conservative 1; Mismatches 6; Indels 0; Gaps 0;

QY      1 LSSATLTSVAKADDFDAQ 18
      : ||||| |||||
Db      3096 ISIGTTLVEVKAIDPDFQ 3113

RESULT 15
US-10-282-122A-51310
; Sequence 51310, Application US/10282122A
; Publication No. US20040029129A1
; GENERAL INFORMATION:
; APPLICANT: Wang, Liangsu
; APPLICANT: Zamudio, Carlos
; APPLICANT: Malone, Cheryl
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Kari
; APPLICANT: Zyskind, Judith
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John
; APPLICANT: Carr, Grant
; APPLICANT: Yamamoto, Robert
; APPLICANT: Forsyth, R.
; APPLICANT: Xu, H.
; TITLE OF INVENTION: Identification of Essential Genes in Microorganisms
; FILE REFERENCE: ELITRA.034A
; CURRENT APPLICATION NUMBER: US/10/282,122A
; CURRENT FILING DATE: 2003-02-20
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/230,335
; PRIOR FILING DATE: 2000-09-06
; PRIOR APPLICATION NUMBER: 60/230,347
; PRIOR FILING DATE: 2000-09-09
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27

```

```

; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/267,636
; PRIOR FILING DATE: 2001-02-09
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 78614
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 51310
; LENGTH: 446
; TYPE: PRT
; ORGANISM: Bordetella pertussis
US-10-282-122A-51310

Query Match          47.8%; Score 44; DB 4; Length 446;
Best Local Similarity 64.3%; Pred. No. 1.3e+02;
Matches 9; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY      5 TTLSAVKADDFDAQ 18
      : |||||
Db      66 TTLRLMADGFDQAQ 79

Search completed: January 25, 2006, 20:10:03
Job time : 61 secs

```

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2006 Compugen Ltd.

OM protein - protein search, using sw model

Run on: January 25, 2006, 19:15:26 ; Search time 5.5 Seconds
(without alignments)
39.378 Million cell updates/sec

Title: US-10-797-821-5
Perfect score: 92
Sequence: 1 LSSATTL SAVKADDFDAQIA 20

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 75621 seqs, 10829074 residues

Total number of hits satisfying chosen parameters: 75621

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published Applications AA New:
1: /cgn2_6/protdata/2/pubpaa/US08_NEW_PUB.pep:
2: /cgn2_6/protdata/2/pubpaa/US06_NEW_PUB.pep:
3: /cgn2_6/protdata/2/pubpaa/US07_NEW_PUB.pep:
4: /cgn2_6/protdata/2/pubpaa/US09_NEW_PUB.pep:
5: /cgn2_6/protdata/2/pubpaa/US10_NEW_PUB.pep:
6: /cgn2_6/protdata/2/pubpaa/US11_NEW_PUB.pep:
7: /cgn2_6/protdata/2/pubpaa/US12_NEW_PUB.pep:
8: /cgn2_6/protdata/2/pubpaa/US13_NEW_PUB.pep:

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	92	100.0	431	7	US-11-052-554A-210 Sequence 210; App
2	39	42.4	122	6	US-10-467-657-6432 Sequence 6432; Ap
3	39	42.4	255	6	US-10-467-657-6432 Sequence 1846; Ap
4	39	42.4	277	7	US-10-454-437-8 Sequence 8; Appli
5	39	42.4	277	7	US-11-055-822-608 Sequence 608; App
6	38	41.3	1060	7	US-11-090-739-120 Sequence 120; App
7	37	40.2	18	6	US-10-981-873-15 Sequence 15; Appli
8	37	40.2	20	6	US-10-981-873-4 Sequence 4; Appli
9	37	40.2	360	7	US-11-169-630-4 Sequence 4; Appli
10	37	40.2	360	7	US-11-169-630-6 Sequence 6; Appli
11	37	40.2	431	6	US-10-485-517-182 Sequence 182; App
12	37	40.2	1107	6	US-10-485-517-145 Sequence 145; App
13	36	39.1	88	6	US-10-467-657-1808 Sequence 1808; Ap
14	36	39.1	182	7	US-11-074-176-218 Sequence 218; App
15	36	39.1	571	6	US-10-793-626-118 Sequence 118; App
16	36	39.1	8595	7	US-11-205-109-15 Sequence 15; Appli
17	35	38.0	162	6	US-10-467-657-6218 Sequence 6218; Ap
18	35	38.0	252	7	US-11-054-515-1467 Sequence 1467; Ap
19	35	38.0	270	6	US-10-793-626-2720 Sequence 2720; Ap
20	35	38.0	338	6	US-10-467-657-4976 Sequence 4976; Ap
21	35	38.0	436	6	US-10-467-657-4550 Sequence 4550; Ap
22	35	38.0	457	7	US-11-112-882-28 Sequence 28; Appli
23	35	38.0	461	7	US-11-075-185-25 Sequence 25; Appli
24	35	38.0	633	6	US-10-453-372-1072 Sequence 1072; Ap
25	35	38.0	633	6	US-10-453-372-1078 Sequence 1078; Ap

26	35	38.0	700	7	US-11-186-284-141	Sequence 141; App
27	35	38.0	733	7	US-11-145-035-42	Sequence 42; Appl
28	35	38.0	853	7	US-11-052-554A-149	Sequence 149; App
29	35	38.0	914	7	US-11-052-554A-160	Sequence 160; App
30	35	38.0	959	6	US-10-453-372-1074	Sequence 1074; Ap
31	35	38.0	1021	6	US-10-453-372-1076	Sequence 1076; Ap
32	35	38.0	1025	6	US-10-453-372-1068	Sequence 1068; Ap
33	35	38.0	1025	6	US-10-453-372-1070	Sequence 1070; Ap
34	35	38.0	1063	6	US-10-453-372-1066	Sequence 1066; Ap
35	35	38.0	1188	7	US-11-143-984A-27	Sequence 27; Appl
36	35	38.0	2828	7	US-11-080-991-54	Sequence 54; Appl
37	35	38.0	2828	7	US-11-186-284-49	Sequence 49; Appl
38	35	38.0	3003	6	US-10-453-372-1080	Sequence 1080; Ap
39	35	38.0	3361	6	US-10-453-372-1082	Sequence 1082; Ap
40	34.5	37.5	534	6	US-10-454-437-16	Sequence 16; Appl
41	34.5	37.5	570	6	US-10-454-437-14	Sequence 14; Appl
42	34.5	37.5	897	6	US-10-336-263A-58	Sequence 58; Appl
43	34.5	37.5	1059	6	US-10-336-263A-54	Sequence 54; Appl
44	34.5	37.5	1059	6	US-10-336-263A-56	Sequence 56; Appl
45	34.5	37.5	3483	6	US-10-453-372-40	Sequence 40; Appl

ALIGNMENTS

RESULT 1
US-11-052-554A-210
; Sequence 210, Application US/11052554A
; Publication No. US2005028866A1
; GENERAL INFORMATION:
; APPLICANT: Sachdeva, et al.
; TITLE OF INVENTION: COMPUTATIONAL METHOD FOR IDENTIFYING ADHESIN AND ADHESIN-LIKE
; FILE REFERENCE: 30853/40359A
; CURRENT APPLICATION NUMBER: US/11/052,554A
; CURRENT FILING DATE: 2005-02-07
; PRIOR APPLICATION NUMBER: US 60/589,227
; PRIOR FILING DATE: 2004-07-20
; PRIOR APPLICATION NUMBER: IN 173/DEL/2004
; PRIOR FILING DATE: 2004-02-06
; NUMBER OF SEQ ID NOS: 763
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 210
; LENGTH: 431
; TYPE: PRT
; ORGANISM: Streptococcus mutans UA159
US-11-052-554A-210

Query Match 100.0%; Score 92; DB 7; Length 431;
Best Local Similarity 100.0%; Pred. No. 2.7e-08;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 LSSATTL SAVKADDFDAQIA 20
DB 16 LSSATTL SAVKADDFDAQIA 35

RESULT 2
US-10-467-657-6432
; Sequence 6432, Application US/10467657
; Publication No. US20050260581A1
; GENERAL INFORMATION:
; APPLICANT: CHIRON Spa
; APPLICANT: FONTANA Maria Rita
; APPLICANT: PIZZA Mariagrazia
; APPLICANT: MASIGNANI Vega
; APPLICANT: MONACI Elisabetta
; TITLE OF INVENTION: GONOCOCCAL PROTEINS AND NUCLEIC ACIDS
; FILE REFERENCE:
; CURRENT APPLICATION NUMBER: US/10/467,657
; CURRENT FILING DATE: 2003-08-11
; PRIOR APPLICATION NUMBER: GB-0103424.8
; PRIOR FILING DATE: 2001-02-12

```
; NUMBER OF SEQ ID NOS: 9218
; SOFTWARE: SeqWin99, version 1.04
; SEQ ID NO 6432
; LENGTH: 122
; TYPE: PRT
; ORGANISM: Neisseria gonorrhoeae
US-10-467-657-6432

Query Match          42.4%; Score 39; DB 6; Length 122;
Best Local Similarity 58.3%; Pred. No. 5.6;
Matches 7; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 9 AVKADDFDAQIA 20
Db 54 SVKDDDFDKAVA 65

RESULT 3
US-10-467-657-1846
; Sequence 1846, Application US/10467657
; Publication No. US20050260591A1
; GENERAL INFORMATION:
; APPLICANT: CHIRON SPA
; APPLICANT: FONTANA Maria Rita
; APPLICANT: PIZZA Mariagrazia
; APPLICANT: MASIGNANI Vega
; APPLICANT: MONACI Elisabetta
; TITLE OF INVENTION: GONOCOCCAL PROTEINS AND NUCLEIC ACIDS
; FILE REFERENCE:
; CURRENT APPLICATION NUMBER: US/10/467,657
; CURRENT FILING DATE: 2003-08-11
; PRIOR APPLICATION NUMBER: GB-0103424.8
; PRIOR FILING DATE: 2001-02-12
; NUMBER OF SEQ ID NOS: 9218
; SOFTWARE: SeqWin99, version 1.04
; SEQ ID NO 1846
; LENGTH: 255
; TYPE: PRT
; ORGANISM: Neisseria gonorrhoeae
US-10-467-657-1846

Query Match          42.4%; Score 39; DB 6; Length 255;
Best Local Similarity 52.9%; Pred. No. 13;
Matches 9; Conservative 2; Mismatches 6; Indels 0; Gaps 0;

QY 4 ATTL5AVKADDFDAQIA 20
Db 22 AVVF5AAVAVDFDARAA 38

RESULT 4
US-10-454-437-8
; Sequence 8, Application US/10454437
; Publication No. US2005027115A1
; GENERAL INFORMATION:
; APPLICANT: Pompejus, Markus
; APPLICANT: Kroger, Burkhard
; APPLICANT: Schroder, Hartwig
; APPLICANT: Zelder, Oskar
; APPLICANT: Haberhauser, Gregor
; TITLE OF INVENTION: INVOLVED IN HOMEOSTASIS AND ADAPTATION
; FILE REFERENCE: BGI-128FCFN
; CURRENT APPLICATION NUMBER: US/10/454,437
; CURRENT FILING DATE: 2003-06-13
; PRIOR APPLICATION NUMBER: US 60/141031
; PRIOR FILING DATE: 1999-06-25
; PRIOR APPLICATION NUMBER: DE 19931636.8
; PRIOR FILING DATE: 1999-07-08
; PRIOR APPLICATION NUMBER: DE 19932125.6
; PRIOR FILING DATE: 1999-07-09
; PRIOR APPLICATION NUMBER: DE 19932126.4
; PRIOR FILING DATE: 1999-07-09

; NUMBER OF SEQ ID NOS: 1158
; Remaining Prior Application data removed - See File Wrapper or PALM.
; SEQ ID NO 608
; LENGTH: 277
; TYPE: PRT
; ORGANISM: Corynebacterium glutamicum
US-11-055-822-608

Query Match          42.4%; Score 39; DB 7; Length 277;
Best Local Similarity 58.3%; Pred. No. 15;
Matches 7; Conservative 4; Mismatches 1; Indels 0; Gaps 0;

QY 9 AVKADDFDAQIA 20
Db 86 AIQADDDAQA 97

RESULT 5
US-11-055-822-608
; Sequence 608, Application US/11055822
; Publication No. US20050260707A1
; GENERAL INFORMATION:
; APPLICANT: Pompejus, Markus
; APPLICANT: Kroger, Burkhard
; APPLICANT: Schroder, Hartwig
; APPLICANT: Zelder, Oskar
; APPLICANT: Haberhauser, Gregor
; TITLE OF INVENTION: CORYNEBACTERIUM GLUTAMICUM GENES ENCODING
; FILE REFERENCE: BGI-121PCPN
; CURRENT APPLICATION NUMBER: US/11/055,822
; CURRENT FILING DATE: 2005-02-11
; PRIOR APPLICATION NUMBER: 09/606,740
; PRIOR FILING DATE: 2000-06-23
; PRIOR APPLICATION NUMBER: 60/141,031
; PRIOR FILING DATE: 1999-06-25
; PRIOR APPLICATION NUMBER: 60/142,101
; PRIOR FILING DATE: 1999-07-02
; PRIOR APPLICATION NUMBER: 60/148,613
; PRIOR FILING DATE: 1999-08-12
; PRIOR APPLICATION NUMBER: 60/187,970
; PRIOR FILING DATE: 2000-03-09
; PRIOR APPLICATION NUMBER: DE 19930476.9
; PRIOR FILING DATE: 1999-07-01
; PRIOR APPLICATION NUMBER: DE 19931415.2
; PRIOR FILING DATE: 1999-07-08
; PRIOR APPLICATION NUMBER: DE 19931418.7
; PRIOR FILING DATE: 1999-07-08
; PRIOR APPLICATION NUMBER: DE 19931419.5
; PRIOR FILING DATE: 1999-07-08
; PRIOR APPLICATION NUMBER: DE 19931420.9
; PRIOR FILING DATE: 1999-07-08
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 1158
; SEQ ID NO 608
; LENGTH: 277
; TYPE: PRT
; ORGANISM: Corynebacterium glutamicum
US-11-055-822-608
```

```

Best Local Similarity 58.3%; Pred. No. 15;
Matches 7; Conservative 4; Mismatches 1; Indels 0; Gaps 0;

QY 9 AVKADDFDAQIA 20
Db 86 AIQAEDDAQVA 97

RESULT 6
US-11-090-739-120
; Sequence 120, Application US/11090739
; Publication No. US20050260639A1
; GENERAL INFORMATION:
; APPLICANT: NAKAMURA, Yusuke
; APPLICANT: KATAGIRI, Toyomasa
; APPLICANT: NAKAGAWA, Hirowaki
; TITLE OF INVENTION: METHOD FOR DIAGNOSING PANCREATIC CANCER
; FILE REFERENCE: 082368-003600US
; CURRENT APPLICATION NUMBER: US/11/090,739
; CURRENT FILING DATE: 2005-03-24
; PRIOR APPLICATION NUMBER: PCT/JP2003/011817
; PRIOR FILING DATE: 2003-09-17
; PRIOR APPLICATION NUMBER: US 60/555,809
; PRIOR FILING DATE: 2004-03-24
; PRIOR APPLICATION NUMBER: US 60/450,889
; PRIOR FILING DATE: 2003-02-28
; PRIOR APPLICATION NUMBER: US 60/414,872
; PRIOR FILING DATE: 2002-09-30
; NUMBER OF SEQ ID NOS: 153
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 120
; LENGTH: 1060
; TYPE: PRT
; ORGANISM: Homo sapiens
US-11-090-739-120

Query Match 41.3%; Score 38; DB 7; Length 1060;
Best Local Similarity 50.0%; Pred. No. 1e+02;
Matches 7; Conservative 4; Mismatches 3; Indels 0; Gaps 0;

QY 6 TLSAVKADDFDAQI 19
Db 735 TVSQVAEDFDSGV 748

RESULT 7
US-10-981-873-15
; Sequence 15, Application US/10981873
; Publication No. US2005025080A1
; GENERAL INFORMATION:
; APPLICANT: Walensky, Loren D.
; APPLICANT: Kormeyer, Stanley J.
; TITLE OF INVENTION: STABILIZED ALPHA HELICAL PEPTIDES AND
; FILE REFERENCE: 00005.001198
; CURRENT APPLICATION NUMBER: US/10/981,873
; CURRENT FILING DATE: 2004-11-05
; PRIOR APPLICATION NUMBER: US 60/517,848
; PRIOR FILING DATE: 2003-11-05
; PRIOR APPLICATION NUMBER: US 60/591,548
; PRIOR FILING DATE: 2004-07-27
; NUMBER OF SEQ ID NOS: 117
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 15
; LENGTH: 18
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Naturally occurring peptide
US-10-981-873-15

Query Match 40.2%; Score 37; DB 6; Length 18;
Best Local Similarity 70.0%; Pred. No. 43;
Matches 7; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 11 KADDFDAQIA 20
Db 54 KGDSYDAQIA 63

Best Local Similarity 53.6%; Pred. No. 1.3;
Matches 7; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 9 AVKADDFDAQI 19
Db 7 AAMCDDFDAQM 17

RESULT 8
US-10-981-873-4
; Sequence 4, Application US/10981873
; Publication No. US20050250680A1
; GENERAL INFORMATION:
; APPLICANT: Walensky, Loren D.
; APPLICANT: Kormeyer, Stanley J.
; APPLICANT: Verdine, Gregory
; TITLE OF INVENTION: STABILIZED ALPHA HELICAL PEPTIDES AND
; FILE REFERENCE: 00530-124001
; CURRENT APPLICATION NUMBER: US/10/981,873
; CURRENT FILING DATE: 2004-11-05
; PRIOR APPLICATION NUMBER: US 60/517,848
; PRIOR FILING DATE: 2003-11-05
; PRIOR APPLICATION NUMBER: US 60/591,548
; PRIOR FILING DATE: 2004-07-27
; NUMBER OF SEQ ID NOS: 117
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 4
; LENGTH: 20
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Naturally occurring peptide
US-10-981-873-4

Query Match 40.2%; Score 37; DB 6; Length 20;
Best Local Similarity 63.6%; Pred. No. 1.5;
Matches 7; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 9 AVKADDFDAQI 19
Db 10 AAMCDDFDAQM 20

RESULT 9
US-11-169-630-4
; Sequence 4, Application US/11169630
; Publication No. US20050266517A1
; GENERAL INFORMATION:
; APPLICANT: KYOWA HAKKO KOGYO CO., LTD
; TITLE OF INVENTION: Novel Transaldolase
; FILE REFERENCE: 00005.001198
; CURRENT APPLICATION NUMBER: US/11/169,630
; CURRENT FILING DATE: 2005-06-30
; PRIOR APPLICATION NUMBER: US/10/088,594
; PRIOR FILING DATE: 2002-03-21
; PRIOR APPLICATION NUMBER: JP 99/266548
; PRIOR FILING DATE: 1999-09-21
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 4
; LENGTH: 360
; TYPE: PRT
; ORGANISM: Corynebacterium glutamicum ATCC31388
US-11-169-630-4

Query Match 40.2%; Score 37; DB 7; Length 360;
Best Local Similarity 70.0%; Pred. No. 43;
Matches 7; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 11 KADDFDAQIA 20
Db 54 KGDSYDAQIA 63

```

```

; APPLICANT: Biosynexus Incorporated
; APPLICANT: Foster, James
; APPLICANT: Mond, Simon
; TITLE OF INVENTION: Antigenic Polypeptides
; FILE REFERENCE: P100629W0
; CURRENT APPLICATION NUMBER: US/10/485,517
; CURRENT FILING DATE: 2004-02-02
; PRIOR APPLICATION NUMBER: GB 0118825.9
; PRIOR FILING DATE: 2001-08-02
; PRIOR APPLICATION NUMBER: GB 0200349.9
; PRIOR FILING DATE: 2002-01-09
; NUMBER OF SEQ ID NOS: 424
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 145
; LENGTH: 1107
; TYPE: PRT
; ORGANISM: Staphylococcus aureus
US-10-485-517-145

Query Match      40.2%; Score 37; DB 6; Length 1107;
Best Local Similarity 50.0%; Pred. No. 1.6e+02;
Matches 9; Conservative 3; Mismatches 6; Indels 0; Gaps 0;

QY      1 LSSATTLSAVKADDFDAQ 18
      :|||:|||:|:|
Db      477 IDSETTRAVKQDLTEAQ 494

RESULT 13
US-10-467-657-1808
; Sequence 1808, Application US/10467657
; Publication No. US20050260581A1
; GENERAL INFORMATION:
; APPLICANT: CHIRON SpA
; APPLICANT: FONTANA Maria Rita
; APPLICANT: PIZZA Mariagrazia
; APPLICANT: MASIGNANI Vega
; APPLICANT: MONACI Elisabetta
; TITLE OF INVENTION: GONOCOCCAL PROTEINS AND NUCLEIC ACIDS
; FILE REFERENCE:
; CURRENT APPLICATION NUMBER: US/10/467,657
; CURRENT FILING DATE: 2003-08-11
; PRIOR APPLICATION NUMBER: GB-0103424.8
; PRIOR FILING DATE: 2001-02-12
; NUMBER OF SEQ ID NOS: 9218
; SOFTWARE: Seqwin99, version 1.04
; SEQ ID NO 1808
; LENGTH: 88
; TYPE: PRT
; ORGANISM: Neisseria gonorrhoeae
US-10-467-657-1808

Query Match      39.1%; Score 36; DB 6; Length 88;
Best Local Similarity 40.0%; Pred. No. 12;
Matches 8; Conservative 6; Mismatches 6; Indels 0; Gaps 0;

QY      1 LSSATTLSAVKADDFDAQIA 20
      :|||:|||:|:|
Db      25 MTAATVLAALSSPVFAAQTA 44

RESULT 14
US-11-074-176-218
; Sequence 218, Application US/11074176
; Publication No. US20050250135A1
; GENERAL INFORMATION:
; APPLICANT: Klaenhammer, Todd R.
; APPLICANT: Russell, William M.
; APPLICANT: Altermann, Eric
; APPLICANT: McAuliffe, Olivia
; APPLICANT: Perill, Andrea Azcarate
; TITLE OF INVENTION: Nucleic Acid Sequences Encoding
; TITLE OF INVENTION: Stress-Related Proteins and Uses Therefore

US-11-169-630-6
; Sequence 6, Application US/11169630
; Publication No. US20050266517A1
; GENERAL INFORMATION:
; APPLICANT: KYOWA HAKKO KOGYO CO., LTD
; TITLE OF INVENTION: Novel Transaldolase
; FILE REFERENCE: 00005.001198
; CURRENT APPLICATION NUMBER: US/11/169,630
; CURRENT FILING DATE: 2005-06-30
; PRIOR APPLICATION NUMBER: US/10/088,594
; PRIOR FILING DATE: 2002-03-21
; PRIOR APPLICATION NUMBER: JP 99/266548
; PRIOR FILING DATE: 1999-09-21
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 6
; LENGTH: 360
; TYPE: PRT
; ORGANISM: Corynebacterium glutamicum ATCC31388
US-11-169-630-6

Query Match      40.2%; Score 37; DB 7; Length 360;
Best Local Similarity 70.0%; Pred. No. 43;
Matches 7; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY      11 KADDFDAQIA 20
      |||:-|||
Db      54 KGDSDYDAQIA 63

RESULT 11
US-10-485-517-182
; Sequence 182, Application US/10485517
; Publication No. US20050256299A1
; GENERAL INFORMATION:
; APPLICANT: University of Sheffield
; APPLICANT: Biosynexus Incorporated
; APPLICANT: Foster, James
; APPLICANT: Mond, Simon
; TITLE OF INVENTION: Antigenic Polypeptides
; FILE REFERENCE: P100629W0
; CURRENT APPLICATION NUMBER: US/10/485,517
; CURRENT FILING DATE: 2004-02-02
; PRIOR APPLICATION NUMBER: GB 0118825.9
; PRIOR FILING DATE: 2001-08-02
; PRIOR APPLICATION NUMBER: GB 0200349.9
; PRIOR FILING DATE: 2002-01-09
; NUMBER OF SEQ ID NOS: 424
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 182
; LENGTH: 431
; TYPE: PRT
; ORGANISM: Staphylococcus aureus
US-10-485-517-182

Query Match      40.2%; Score 37; DB 6; Length 431;
Best Local Similarity 77.8%; Pred. No. 53;
Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY      12 ADFFDAQIA 20
      |||:-|||
Db      240 ADEFDVQIA 248

RESULT 12
US-10-485-517-145
; Sequence 145, Application US/10485517
; Publication No. US20050256299A1
; GENERAL INFORMATION:
; APPLICANT: University of Sheffield
```

```

; FILE REFERENCE: 5051-694
; CURRENT APPLICATION NUMBER: US/11/074,176
; CURRENT FILING DATE: 2005-03-07
; PRIOR APPLICATION NUMBER: 60/551,161
; PRIOR FILING DATE: 2004-03-08
; NUMBER OF SEQ ID NOS: 381
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 218
; LENGTH: 182
; TYPE: PRT
; ORGANISM: Lactobacillus acidophilus
US-11-074-176-218

Query Match      39.1%; Score 36; DB 7; Length 182;
Best Local Similarity 50.0%; Pred. No. 29;
Matches 7; Conservative 3; Mismatches 4; Indels 0; Gaps 0;

QY      1 LSSATTLSAVKADD 14
Db      108 IASGTATTAVKLD 121

RESULT 15
US-10-793-626-118
; Sequence 118, Application US/10793626
; Publication No. US20050255478A1
; GENERAL INFORMATION:
; APPLICANT: KIMMERLY, WILLIAM JOHN
; TITLE OF INVENTION: STAPHYLOCOCCUS EPIDERMIDIS NUCLEIC ACIDS AND PROTEINS
; FILE REFERENCE: PUB480US
; CURRENT APPLICATION NUMBER: US/10/793,626
; CURRENT FILING DATE: 2004-03-04
; PRIOR APPLICATION NUMBER: 60/164,258
; PRIOR FILING DATE: 1999-11-09
; NUMBER OF SEQ ID NOS: 4472
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 118
; LENGTH: 571
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: synthetic
; OTHER INFORMATION: amino acid sequence
US-10-793-626-118

Query Match      39.1%; Score 36; DB 6; Length 571;
Best Local Similarity 66.7%; Pred. No. 1.1e+02;
Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY      12 ADFDAQIA 20
Db      240 ADEFVQVA 248

Search completed: January 25, 2006, 20:11:08
Job time : 6.5 secs

```

THIS PAGE BLANK (USPTO)

GenCore version 5.1.1.6
Copyright (c) 1993 - 2006 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: January 25, 2006, 18:48:24 ; Search time 68.2 Seconds
(without alignments)
128.850 Million cell updates/sec

Title: US-10-797-821-6
Perfect score: 94
Sequence: 1 QIASQDSKINLTAQQAAQ 20

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 2443163 seqs, 439378781 residues

Total number of hits satisfying chosen parameters: 2443163

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : A_Geneseq 21:*

- 1: geneseqp1980s:*
- 2: geneseqp1990s:*
- 3: geneseqp2000s:*
- 4: geneseqp2001s:*
- 5: geneseqp2002s:*
- 6: geneseqp2003as:*
- 7: geneseqp2003bs:*
- 8: geneseqp2004s:*
- 9: geneseqp2005s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	94	100.0	20	7	ADD93626 Streptoco
2	94	100.0	20	9	ADK37249 Streptoco
3	94	100.0	431	7	ADD93649 Streptoco
4	94	100.0	431	7	ADD93650 Streptoco
5	94	100.0	431	7	ADD93653 Streptoco
6	94	100.0	431	9	ADK37272 Streptoco
7	94	100.0	431	9	ADK37273 Streptoco
8	94	100.0	431	9	ADK37276 Streptoco
9	94	100.0	431	9	AEb91500 Microbial
10	94	100.0	432	7	ADD93651 Streptoco
11	94	100.0	432	7	ADD93652 Streptoco
12	94	100.0	432	9	ADK37274 Streptoco
13	94	100.0	432	9	ADK37275 Streptoco
14	77	81.9	20	7	ADD93627 Streptoco
15	77	81.9	20	9	ADK37250 Streptoco
16	74	78.7	210	2	AAY22579 Bacterial
17	74	78.7	392	6	ABU02747 S. pneumo
18	74	78.7	392	8	ADK47859 Streptoco
19	74	78.7	392	8	ADT50327 S.pneumon
20	74	78.7	392	8	ADT50226 S.pneumon
21	74	78.7	392	8	ADT50165 S.pneumon
22	74	78.7	399	8	ADR94595 Novel S.
23	74	78.7	399	9	AEA58465 Streptoco
24	62	66.0	398	5	ABP25919 Streptoco

ALIGNMENTS

RESULT 1

ADD93626

ID ADD93626 standard; peptide; 20 AA.

XX

AC ADD93626;

XX

DT 29-JAN-2004 (first entry)

XX

DE Streptococcus mutans glucan binding protein-B peptide fragment.

XX

KW Glucan binding protein-B; GbpB; vaccine; anticaries; epitope; immunogen.

XX

OS Streptococcus mutans.

XX

PN WO2003075845-A2.

XX

PD 18-SEP-2003.

XX

PF 07-MAR-2003; 2003WO-US006962.

XX

PR 07-MAR-2002; 2002US-0363209P.

PR

08-AUG-2002; 2002US-0402483P.

XX

(FORS-) FORSYTH INST.

PA

Smith DJ, Taubman MA;

XX

WPI; 2003-845091/78.

XX

Composition useful as vaccines for dental caries comprises a fragment of a glucan binding protein-B binding to a major histocompatibility complex class II protein.

XX

PS Claim 11; Page 10; 49pp; English.

XX

The present sequence is that of a peptide comprising amino acid residues 33-52 of the glucan binding protein-B (GbpB) of Streptococcus mutans strain SK32 ADD93649. The peptide binds to a major histocompatibility complex (MHC) class II protein. It was identified as a potential B cell epitope using a matrix-based algorithm for epitope prediction, which was used to search the primary amino acid sequence of GbpB for known MHC class II binding motifs. The peptide can be used in immunogenic compositions and subunit vaccines for dental caries. These comprise MHC class II protein-binding GbpB peptides covalently linked with peptide subunits (preferably from the catalytic domain) of a glucosyltransferase. The compositions are used in a claimed method of eliciting production of an antibody in a mammal. Dieptopic or multieptopic polypeptides can be

Adr83884 S. pyogen
Aeb91542 Microbial
Abp29684 Streptoco
Adu69524 S agalact
Adv88392 Streptoco
Adv81808 Streptoco
Adv78645 Streptoco
Abg75182 M licheni
Abbs7775 Drosophil
Abr40385 Fruitfly
Ads96412 Drosophil
Adc06779 Fruit fly
Adv16553 E. faecal
Adh88105 Enterococ
Abb65663 Drosophil
Abb70956 Plant ful
Aag91717 C glutami
Abb66194 Drosophil
Aay97820 Pseudomon
Abo66368 Klebsiell
Abu29734 Protein e

25 62 66.0 398 8 ADR83884
26 62 66.0 398 9 AEB91542
27 54 57.4 447 5 ABP29684
28 54 57.4 447 8 ADU69524
29 54 57.4 447 8 ADV88392
30 54 57.4 447 8 ADV81808
31 54 57.4 447 8 ADV78645
32 49 52.1 1050 7 ABG75182
33 48 51.1 325 6 ABB57775
34 48 51.1 325 6 ABR40385
35 48 51.1 325 8 ADS96412
36 48 51.1 478 7 ADC06779
37 47 50.0 482 9 ADV16553
38 47 50.0 497 7 ADH88105
39 46 48.9 199 8 ABB65663
40 46 48.9 199 8 ABB70956
41 45 47.9 371 4 AAG91717
42 44 46.8 341 4 ABB66194
43 44 46.8 489 3 AAY97820
44 44 46.8 523 7 ABO66368
45 44 46.8 524 6 ABU29734

CC prepared synthetically or by recombinant DNA technology. Antibodies
CC raised against MHC class II binding fragments of GbpB can be used in
CC passive immunisation.

SQ Sequence 20 AA;

Query Match 100.0%; Score 94; DB 7; Length 20;
Best Local Similarity 100.0%; Pred. No. 2.9e-08;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 QIASQDSKINNLTAQQAAQ 20
| | | | | | | | | | | | | | | | | | | | | |
Db 1 QIASQDSKINNLTAQQAAQ 20

RESULT 2

ADX37249
ID ADX37249 standard; peptide; 20 AA.
XX
AC ADX37249;
XX
DT 21-APR-2005 (first entry)
XX
DE Streptococcus mutant glucan binding protein B peptide #6.
XX immunogenicity; immune stimulation; glucan binding protein-B;
KW microparticle; major histocompatibility complex; tooth disease.
XX Streptococcus mutans.
OS
PN US2005031633-A1.
XX
PD 10-FEB-2005.
XX
PF 09-MAR-2004; 2004US-00797821.
XX
PR 13-APR-1998; 98US-0081550P.
PR 08-JAN-1999; 99US-0115142P.
PR 12-APR-1999; 99US-00290049.
PR 07-MAR-2002; 2002US-0363309P.
PR 08-AUG-2002; 2002US-0402483P.
PR 07-MAR-2003; 2003US-00383930.
XX
PA (SMIT/) SMITH D J.
PA (TAUB/) TAUBMAN M A.
PI
PI Smith DJ, Taubman MA;
XX
XX WPI; 2005-151644/16.

XX New composition comprising a fragment of a glucan binding protein-B
PT (GbpB) that binds to MHC class II protein, and a biocompatible
PT microparticle, useful for producing an antibody (claimed) for immunizing
PT mammals against dental caries.
XX
PS Claim 4; SEQ ID NO 6; 73pp; English.
XX
XX The invention relates to a composition comprising a fragment of a glucan
CC binding protein-B (GbpB) and a biocompatible microparticle, where the
CC fragment binds to a major histocompatibility complex (MHC) class II
CC protein. The composition is useful for producing an antibody for
CC immunizing mammals against dental caries. This sequence corresponds to a
CC Streptococcus mutans GbpB-derived peptide of the invention.
XX
SQ Sequence 20 AA;

Query Match 100.0%; Score 94; DB 9; Length 20;
Best Local Similarity 100.0%; Pred. No. 2.9e-08;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 QIASQDSKINNLTAQQAAQ 20
| | | | | | | | | | | | | | | | | | | | | |
Db 1 QIASQDSKINNLTAQQAAQ 20

RESULT 3
ADD93649
ID ADD93649 standard; protein; 431 AA.
XX
AC ADD93649;
XX
DT 29-JAN-2004 (first entry)
XX
DE Streptococcus mutans glucan binding protein-B.
KW Glucan binding protein-B; GbpB; vaccine; anticaries; epitope; immunogen.
XX Streptococcus mutans.
OS
FH Key Location/Qualifiers
FT Region 6..25
FT /note= "HLA-binding peptide"
FT Region 16..35
FT /note= "HLA-binding peptide"
FT Region 33..52
FT /note= "HLA-binding peptide"
FT Region 37..56
FT /note= "HLA-binding peptide"
FT Region 48..67
FT /note= "HLA-binding peptide"
FT Region 52..71
FT /note= "HLA-binding peptide"
FT Region 88..107
FT /note= "HLA-binding peptide"
FT Region 113..132
FT /note= "HLA-binding peptide"
FT Region 117..136
FT /note= "HLA-binding peptide"
FT Region 137..156
FT /note= "HLA-binding peptide"
FT Region 174..193
FT /note= "HLA-binding peptide"
FT Region 194..213
FT /note= "HLA-binding peptide"
FT Region 214..233
FT /note= "HLA-binding peptide"
FT Region 248..267
FT /note= "HLA-binding peptide"
FT Region 289..308
FT /note= "HLA-binding peptide"
FT Region 306..325
FT /note= "HLA-binding peptide"
FT Region 311..330
FT /note= "HLA-binding peptide"
FT Region 349..368
FT /note= "HLA-binding peptide"
FT Region 365..384
FT /note= "HLA-binding peptide"
FT Region 383..402
FT /note= "HLA-binding peptide"
FT Region 403..422
FT /note= "HLA-binding peptide"
XX WO2003075845-A2.
XX 18-SEP-2003.
XX
XX 07-MAR-2003; 2003WO-US006962.
XX
XX 07-MAR-2002; 2002US-0363209P.
PR 08-AUG-2002; 2002US-0402483P.
XX
PA (FORS-) FORSYTH INST.
XX Smith DJ, Taubman MA;
XX

DR WPI; 2003-845091/78.
 DR GENBANK; AY046410.
 XX
 PT Composition useful as vaccines for dental caries comprises a fragment of
 PT a glucan binding protein-B binding to a major histocompatibility complex
 PT class II protein.
 XX
 PS Claim 6; Page 7; 49pp; English.
 XX
 CC The present sequence is the protein sequence of the glucan binding
 CC protein-B (GbpB) of Streptococcus mutans strain SK32. The sequence
 CC includes a number of human leucocyte antigen (HLA)-binding epitopes. The
 CC invention provides immunogenic compositions and vaccines for dental
 CC caries. The compositions comprise major histocompatibility complex (MHC)
 CC class II protein-binding GbpB peptides, especially HLA-binding peptides,
 CC covalently linked with peptide subunits of a glucosyltransferase. The
 CC compositions are used in a claimed method of eliciting production of an
 CC antibody in a mammal. Diepitopic or multiepitopic polypeptides can be
 CC prepared synthetically or by recombinant DNA technology. Antibodies
 CC raised against MHC class II binding fragments of GbpB can be used in
 CC passive immunisation.
 XX
 SQ Sequence 431 AA;
 Query Match 100.0%; Score 94; DB 7; Length 431;
 Best Local Similarity 100.0%; Pred. No. 1e-06;
 Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 OY 1 QIASQDSKINNLTAAQQAAQ 20
 |||||
 DB 33 QIASQDSKINNLTAAQQAAQ 52
 |||||
 RESULT 4
 ADD93650
 ID ADD93650 standard; protein; 431 AA.
 XX
 AC ADD93650;
 XX
 DT 29-JAN-2004 (first entry)
 XX
 DE Streptococcus mutans glucan binding protein-B.
 XX
 KW Glucan binding protein-B; GbpB; vaccine; anticaries; epitope; immunogen.
 XX
 OS Streptococcus mutans.
 XX
 FN WO2003075845-A2.
 XX
 PD 18-SEP-2003.
 XX
 PF 07-MAR-2003; 2003WO-US006962.
 XX
 PR 07-MAR-2002; 2002US-0363209P.
 PR 08-AUG-2002; 2002US-0402483P.
 XX
 PA (FORS-) FORSYTH INST.
 XX
 PI Smith DJ, Taubman MA;
 XX
 DR WPI; 2003-845091/78.
 DR GENBANK; AY046411.
 XX
 PT Composition useful as vaccines for dental caries comprises a fragment of
 PT a glucan binding protein-B binding to a major histocompatibility complex
 PT class II protein.
 XX
 PS Claim 5; Page 8; 49pp; English.
 XX
 CC The present sequence is the protein sequence of the glucan binding
 CC protein-B (GbpB) of Streptococcus mutans strain 3V74. The sequence
 CC includes a number of human leucocyte antigen (HLA)-binding epitopes. The
 CC invention provides immunogenic compositions and vaccines for dental
 CC caries. The compositions comprise major histocompatibility complex (MHC)
 CC class II protein-binding GbpB peptides, especially HLA-binding peptides,
 CC covalently linked with peptide subunits of a glucosyltransferase. The
 CC compositions are used in a claimed method of eliciting production of an
 CC antibody in a mammal. Diepitopic or multiepitopic polypeptides can be
 CC prepared synthetically or by recombinant DNA technology. Antibodies
 CC raised against MHC class II binding fragments of GbpB can be used in
 CC passive immunisation.
 XX
 SQ Sequence 431 AA;
 Query Match 100.0%; Score 94; DB 7; Length 431;
 Best Local Similarity 100.0%; Pred. No. 1e-06;
 Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 OY 1 QIASQDSKINNLTAAQQAAQ 20
 |||||
 DB 33 QIASQDSKINNLTAAQQAAQ 52
 |||||
 RESULT 5
 ADD93653
 ID ADD93653 standard; protein; 431 AA.
 XX
 AC ADD93653;
 XX
 DT 29-JAN-2004 (first entry)
 XX
 DE Streptococcus mutans glucan binding protein-B.
 XX
 KW Glucan binding protein-B; GbpB; vaccine; anticaries; epitope; immunogen.
 XX
 OS Streptococcus mutans.
 XX
 FN WO2003075845-A2.
 XX
 PD 18-SEP-2003.
 XX
 PF 07-MAR-2003; 2003WO-US006962.
 XX
 PR 07-MAR-2002; 2002US-0363209P.
 PR 08-AUG-2002; 2002US-0402483P.
 XX
 PA (FORS-) FORSYTH INST.
 XX
 PI Smith DJ, Taubman MA;
 XX
 DR WPI; 2003-845091/78.
 DR GENBANK; AY046414.
 XX
 PT Composition useful as vaccines for dental caries comprises a fragment of
 PT a glucan binding protein-B binding to a major histocompatibility complex
 PT class II protein.
 XX
 PS Claim 5; Page 8-9; 49pp; English.
 XX
 CC The present sequence is the protein sequence of the glucan binding
 CC protein-B (GbpB) of Streptococcus mutans strain 5SM3. The sequence
 CC includes a number of human leucocyte antigen (HLA)-binding epitopes. The
 CC invention provides immunogenic compositions and vaccines for dental
 CC caries. The compositions comprise major histocompatibility complex (MHC)
 CC class II protein-binding GbpB peptides, especially HLA-binding peptides,
 CC covalently linked with peptide subunits of a glucosyltransferase. The
 CC compositions are used in a claimed method of eliciting production of an
 CC antibody in a mammal. Diepitopic or multiepitopic polypeptides can be
 CC prepared synthetically or by recombinant DNA technology. Antibodies
 CC raised against MHC class II binding fragments of GbpB can be used in
 CC passive immunisation.
 XX
 SQ Sequence 431 AA;
 Query Match 100.0%; Score 94; DB 7; Length 431;
 Best Local Similarity 100.0%; Pred. No. 1e-06;
 Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 QIASQDSKINNLTAAQQAAQ 20
 |||||
 Db 33 QIASQDSKINNLTAAQQAAQ 52

RESULT 6
 ADX37272
 ID ADX37272 standard; protein; 431 AA.
 XX
 AC ADX37272;
 XX
 DT 21-APR-2005 (first entry)
 XX
 DE Streptococcus mutant glucan binding protein B variant #1.
 XX
 KW immunogenicity; immune stimulation; glucan binding protein-B;
 KW microparticle; major histocompatibility complex; tooth disease.
 XX
 OS Streptococcus mutans.
 XX
 PN US2005031633-A1.
 XX
 PD 10-FEB-2005.
 XX
 PF 09-MAR-2004; 2004US-00797821.
 XX
 PR 13-APR-1998; 98US-0081550P.
 PR 08-JAN-1999; 99US-0115142P.
 PR 12-APR-1999; 99US-00290049.
 PR 07-MAR-2002; 2002US-0363209P.
 PR 08-AUG-2002; 2002US-0402483P.
 PR 07-MAR-2003; 2003US-00383930.
 XX
 PA (SMIT/) SMITH D J.
 PA (TAUB/) TAUBMAN M A.
 XX
 PI Smith DJ, Taubman MA;
 XX
 DR WPI; 2005-151644/16.
 XX
 PT New composition comprising a fragment of a glucan binding protein-B
 PT (GbpB) that binds to MHC class II protein, and a biocompatible
 PT microparticle, useful for producing an antibody (claimed) for immunizing
 PT mammals against dental caries.
 XX
 PS Claim 3; SEQ ID NO 29; 73pp; English.
 XX
 CC The invention relates to a composition comprising a fragment of a glucan
 CC binding protein-B (GbpB) and a biocompatible microparticle, where the
 CC fragment binds to a major histocompatibility complex (MHC) class II
 CC protein. The composition is useful for producing an antibody for
 CC immunizing mammals against dental caries. This sequence corresponds to a
 CC Streptococcus mutans GbpB protein of the invention.
 XX
 SQ Sequence 431 AA;
 Query Match 100.0%; Score 94; DB 9; Length 431;
 Best Local Similarity 100.0%; Pred. No. 1e-06;
 Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 QIASQDSKINNLTAAQQAAQ 20
 |||||
 Db 33 QIASQDSKINNLTAAQQAAQ 52

RESULT 7
 ADX37273
 ID ADX37273 standard; protein; 431 AA.
 XX
 AC ADX37273;
 XX

DT 21-APR-2005 (first entry)
 XX
 DE Streptococcus mutant glucan binding protein B variant #2.
 XX
 KW immunogenicity; immune stimulation; glucan binding protein-B;
 KW microparticle; major histocompatibility complex; tooth disease.
 XX
 OS Streptococcus mutans.
 XX
 PN US2005031633-A1.
 XX
 PD 10-FEB-2005.
 XX
 PF 09-MAR-2004; 2004US-00797821.
 XX
 PR 13-APR-1998; 98US-0081550P.
 PR 08-JAN-1999; 99US-0115142P.
 PR 12-APR-1999; 99US-00290049.
 PR 07-MAR-2002; 2002US-0363209P.
 PR 08-AUG-2002; 2002US-0402483P.
 PR 07-MAR-2003; 2003US-00383930.
 XX
 PA (SMIT/) SMITH D J.
 PA (TAUB/) TAUBMAN M A.
 XX
 PI Smith DJ, Taubman MA;
 XX
 DR WPI; 2005-151644/16.
 XX
 PT New composition comprising a fragment of a glucan binding protein-B
 PT (GbpB) that binds to MHC class II protein, and a biocompatible
 PT microparticle, useful for producing an antibody (claimed) for immunizing
 PT mammals against dental caries.
 XX
 PS Claim 3; SEQ ID NO 30; 73pp; English.
 XX
 CC The invention relates to a composition comprising a fragment of a glucan
 CC binding protein-B (GbpB) and a biocompatible microparticle, where the
 CC fragment binds to a major histocompatibility complex (MHC) class II
 CC protein. The composition is useful for producing an antibody for
 CC immunizing mammals against dental caries. This sequence corresponds to a
 CC Streptococcus mutans GbpB protein of the invention.
 XX
 SQ Sequence 431 AA;
 Query Match 100.0%; Score 94; DB 9; Length 431;
 Best Local Similarity 100.0%; Pred. No. 1e-06;
 Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 QIASQDSKINNLTAAQQAAQ 20
 |||||
 Db 33 QIASQDSKINNLTAAQQAAQ 52

RESULT 8
 ADX37276
 ID ADX37276 standard; protein; 431 AA.
 XX
 AC ADX37276;
 XX
 DT 21-APR-2005 (first entry)
 XX
 DE Streptococcus mutant glucan binding protein B variant #5.
 XX
 KW immunogenicity; immune stimulation; glucan binding protein-B;
 KW microparticle; major histocompatibility complex; tooth disease.
 XX
 OS Streptococcus mutans.
 XX
 PN US2005031633-A1.
 XX
 PD 10-FEB-2005.
 XX

XX Composition useful as vaccines for dental caries comprises a fragment of
 PT a glucan binding protein-B binding to a major histocompatibility complex
 PT class II protein.

XX Claim 5; Page 8; 49pp; English.

XX The present sequence is the protein sequence of the glucan binding
 CC protein-B (GbpB) of Streptococcus mutans strain 15JP2. The sequence
 CC includes a number of human leucocyte antigen (HLA)-binding epitopes. The
 CC invention provides immunogenic compositions and vaccines for dental
 CC caries. The compositions comprise major histocompatibility complex (MHC)
 CC class II protein-binding GbpB peptides, especially HLA-binding peptides,
 CC covalently linked with peptide subunits of a glucosyltransferase. The
 CC compositions are used in a claimed method of eliciting production of an
 CC antibody in a mammal. Dieptopic or multiepitopic polypeptides can be
 CC prepared synthetically or by recombinant DNA technology. Antibodies
 CC raised against MHC class II binding fragments of GbpB can be used in
 CC passive immunisation.

XX Sequence 432 AA;

Query Match 100.0%; Score 94; DB 7; Length 432;
 Best Local Similarity 100.0%; Pred. No. 1e-06;
 Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 QIASQDSKINNLTAAQQAAQ 20
 |||||
 Db 33 QIASQDSKINNLTAAQQAAQ 52

RESULT 11

ADD93652
 ID ADD93652 standard; protein; 432 AA.

XX
 AC ADD93652;

DT 29-JAN-2004 (first entry)

DE Streptococcus mutans glucan binding protein-B.

XX Glucan binding protein-B; GbpB; vaccine; anticaries; epitope; immunogen.

XX Streptococcus mutans.

XX WO2003075845-A2.

XX 18-SEP-2003.

XX 07-MAR-2003; 2003WO-US006962.

XX 07-MAR-2002; 2002US-0363209P.

PR 08-AUG-2002; 2002US-0402483P.

XX (FORS-) FORSYTH INST.

XX Smith DJ, Taubman MA;

XX WPI; 2003-845091/78.

DR GENBANK; AY046413.

XX Composition useful as vaccines for dental caries comprises a fragment of
 CC a glucan binding protein-B binding to a major histocompatibility complex
 CC class II protein.

XX Claim 5; Page 8; 49pp; English.

XX The present sequence is the protein sequence of the glucan binding
 CC protein-B (GbpB) of Streptococcus mutans strain 3SN1. The sequence
 CC includes a number of human leucocyte antigen (HLA)-binding epitopes. The
 CC invention provides immunogenic compositions and vaccines for dental
 CC caries. The compositions comprise major histocompatibility complex (MHC)
 CC class II protein-binding GbpB peptides, especially HLA-binding peptides,

CC covalently linked with peptide subunits of a glucosyltransferase. The
 CC compositions are used in a claimed method of eliciting production of an
 CC antibody in a mammal. Dieptopic or multiepitopic polypeptides can be
 CC prepared synthetically or by recombinant DNA technology. Antibodies
 CC raised against MHC class II binding fragments of GbpB can be used in
 CC passive immunisation.

XX Sequence 432 AA;

Query Match 100.0%; Score 94; DB 7; Length 432;
 Best Local Similarity 100.0%; Pred. No. 1e-06;
 Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 QIASQDSKINNLTAAQQAAQ 20
 |||||
 Db 33 QIASQDSKINNLTAAQQAAQ 52

RESULT 12

ADX37274
 ID ADX37274 standard; protein; 432 AA.

XX
 AC ADX37274;

DT 21-APR-2005 (first entry)

XX Streptococcus mutant glucan binding protein B variant #3.

XX immunogenicity; immune stimulation; glucan binding protein-B;
 KW microparticle; major histocompatibility complex; tooth disease.

XX Streptococcus mutans.

XX US2005031633-A1.

XX 10-FEB-2005.

XX 09-MAR-2004; 2004US-00797821.

XX 13-APR-1998; 98US-0081550P.

PR 08-JAN-1999; 99US-0115142P.

PR 12-APR-1999; 99US-0029004P.

PR 07-MAR-2002; 2002US-0363209P.

PR 08-AUG-2002; 2002US-0402483P.

PR 07-MAR-2003; 2003US-00383930.

XX (SMIT/) SMITH D J.

PA (TAUB/) TAUBMAN M A.

XX Smith DJ, Taubman MA;

XX WPI; 2005-151644/16.

XX New composition comprising a fragment of a glucan binding protein-B
 PT (GbpB) that binds to MHC class II protein, and a biocompatible
 PT microparticle, useful for producing an antibody (claimed) for immunizing
 PT mammals against dental caries.

XX Claim 3; SEQ ID NO 31; 73pp; English.

XX The invention relates to a composition comprising a fragment of a glucan
 CC binding protein-B (GbpB) and a biocompatible microparticle, where the
 CC fragment binds to a major histocompatibility complex (MHC) class II
 CC protein. The composition is useful for producing an antibody for
 CC immunizing mammals against dental caries. This sequence corresponds to a
 CC Streptococcus mutans GbpB protein of the invention.

XX Sequence 432 AA;

Query Match 100.0%; Score 94; DB 9; Length 432;
 Best Local Similarity 100.0%; Pred. No. 1e-06;
 Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 QIASQDSKINNLTAQOQAAQ 20
 Db 33 QIASQDSKINNLTAQOQAAQ 52

RESULT 13
 ADX37275
 ID ADX37275 standard; protein; 432 AA.
 XX AC ADX37275;
 XX DT 21-APR-2005 (first entry)
 XX DE Streptococcus mutant glucan binding protein B variant #4.
 KW immunogenicity; immune stimulation; glucan binding protein-B;
 KW microparticle; major histocompatibility complex; tooth disease.
 XX OS Streptococcus mutans.
 XX PN US2005031633-A1.
 PD 10-FEB-2005.
 XX PF 09-MAR-2004; 2004US-00797821.
 XX PR 13-APR-1998; 98US-0081550P.
 PR 08-JAN-1999; 99US-0115142P.
 PR 12-APR-1999; 99US-00290049.
 PR 07-MAR-2002; 2002US-0363209P.
 PR 08-AUG-2002; 2002US-0402483P.
 PR 07-MAR-2003; 2003US-00383930.
 XX (SMIT/) SMITH D J.
 PA (TAUB/) TAUBMAN M A.
 PI Smith DJ, Taubman MA;
 XX WPI; 2005-151644/16.
 DR New composition comprising a fragment of a glucan binding protein-B
 PT (GbpB) that binds to MHC class II protein, and a biocompatible
 PT microparticle, useful for producing an antibody (claimed) for immunizing
 PT mammals against dental caries.
 XX Claim 3; SEQ ID NO 32; 73pp; English.
 CC The invention relates to a composition comprising a fragment of a glucan
 CC binding protein-B (GbpB) and a biocompatible microparticle, where the
 CC fragment binds to a major histocompatibility complex (MHC) class II
 CC protein. The composition is useful for producing an antibody for
 CC immunizing mammals against dental caries. This sequence corresponds to a
 CC Streptococcus mutans GbpB protein of the invention.
 XX Sequence 432 AA;
 SQ Query Match 100.0%; Score 94; DB 9; Length 432;
 Best Local Similarity 100.0%; Pred. No. 1e-06;
 Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 QIASQDSKINNLTAQOQAAQ 20
 Db 33 QIASQDSKINNLTAQOQAAQ 52

RESULT 14
 ADD93627
 ID ADD93627 standard; peptide; 20 AA.
 XX AC ADD93627;
 XX DT 29-JAN-2004 (first entry)
 XX

DE Streptococcus mutans glucan binding protein-B peptide fragment.
 XX Glucan binding protein-B; GbpB; vaccine; anticaries; epitope; immunogen.
 XX Streptococcus mutans.
 XX WO2003075845-A2.
 XX PD 18-SEP-2003.
 XX PF 07-MAR-2003; 2003WO-US006962.
 XX PR 07-MAR-2002; 2002US-0363209P.
 PR 08-AUG-2002; 2002US-0402483P.
 XX (FORS-) FORSYTH INST.
 PA Smith DJ, Taubman MA;
 XX WPI; 2003-845091/78.
 DR Composition useful as vaccines for dental caries comprises a fragment of
 PT a glucan binding protein-B binding to a major histocompatibility complex
 PT class II protein.
 XX Claim 11; Page 10; 49pp; English.
 XX The present sequence is that of a peptide comprising amino acid residues
 CC 37-56 of the glucan binding protein-B (GbpB) of Streptococcus mutans
 CC strain SK32 ADD93649. The peptide binds to a major histocompatibility
 CC complex (MHC) class II protein. It was identified as a potential B cell
 CC epitope using a matrix-based algorithm for epitope prediction, which was
 CC used to search the primary amino acid sequence of GbpB for known MHC
 CC class II binding motifs. The peptide can be used in immunogenic
 CC compositions and subunit vaccines for dental caries. These comprise MHC
 CC class II protein-binding GbpB peptides covalently linked with peptide
 CC subunits (preferably from the catalytic domain) of a glucosyltransferase.
 CC The compositions are used in a claimed method of eliciting production of
 CC an antibody in a mammal. Dieptopic or multiepitopic polypeptides can be
 CC prepared synthetically or by recombinant DNA technology. Antibodies
 CC raised against MHC class II binding fragments of GbpB can be used in
 CC passive immunisation.
 XX Sequence 20 AA;
 SQ Query Match 81.9%; Score 77; DB 7; Length 20;
 Best Local Similarity 100.0%; Pred. No. 1.8e-05;
 Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 5 QDSKINNLTAQOQAAQ 20
 Db 1 QDSKINNLTAQOQAAQ 16

RESULT 15
 ADX37250
 ID ADX37250 standard; peptide; 20 AA.
 XX AC ADX37250;
 XX DT 21-APR-2005 (first entry)
 XX DE Streptococcus mutant glucan binding protein B peptide #7.
 KW immunogenicity; immune stimulation; glucan binding protein-B;
 KW microparticle; major histocompatibility complex; tooth disease.
 XX OS Streptococcus mutans.
 XX PN US2005031633-A1.
 XX PD 10-FEB-2005.
 XX

PF 09-MAR-2004; 2004US-00797821.
XX
PR 13-APR-1998; 98US-0081550P.
PR 08-JAN-1999; 99US-0115142P.
PR 12-APR-1999; 99US-00290049.
PR 07-MAR-2002; 2002US-0363209P.
PR 08-AUG-2002; 2002US-0402483P.
PR 07-MAR-2003; 2003US-00383930.
XX
XX (SMIT/) SMITH D J.
PA (TAUB/) TAUBMAN M A.
XX
PI Smith DJ, Taubman MA;
XX
DR WPI; 2005-151644/16.
XX
XX New composition comprising a fragment of a glucan binding protein-B
PT (GbpB) that binds to MHC class II protein, and a biocompatible
PT microparticle, useful for producing an antibody (claimed) for immunizing
PT mammals against dental caries.
XX
PS Claim 4; SEQ ID NO 7; 73pp; English.
XX
CC The invention relates to a composition comprising a fragment of a glucan
CC binding protein-B (GbpB) and a biocompatible microparticle, where the
CC fragment binds to a major histocompatibility complex (MHC), class II
CC protein. The composition is useful for producing an antibody for
CC immunizing mammals against dental caries. This sequence corresponds to a
CC Streptococcus mutans GbpB-derived peptide of the invention.
XX
SQ Sequence 20 AA;

Query Match 81.9%; Score 77; DB 9; Length 20;
Best Local Similarity 100.0%; Pred. No. 1.8e-05;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 5 QDSKINNLTAAQQAAQ 20
Db 1 QDSKINNLTAAQQAAQ 16

Search completed: January 25, 2006, 19:01:43
Job time : 68.2 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2006 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: January 25, 2006, 18:48:55 ; Search time 9.1 Seconds
(without alignments)
211.465 Million cell updates/sec

Title: US-10-797-821-6

Perfect score: 94

Sequence: 1 QIASQDSKINLTAQQQAAQ 20

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 283416 seqs, 96216763 residues

Total number of hits satisfying chosen parameters: 283416

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

1: Pir1.*
2: Pir2.*
3: Pir3.*
4: Pir4.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
1	74	78.7	392	2 G95258	secreted 45 kd pro
2	74	78.7	392	2 B98124	general stress
3	47	50.0	507	2 S05542	hypothetical prote
4	46	48.9	238	2 D75477	hypothetical prote
5	46	48.9	503	2 AB2734	NADH dehydrogenase
6	46	48.9	503	2 D97515	NADH dehydrogenase
7	45	47.9	526	2 E88940	protein C05B4_9 [i
8	44	46.8	308	2 AF1990	hypothetical prote
9	44	46.8	505	2 S15194	sucrose porin' scry
10	44	46.8	947	2 T23107	hypothetical prote
11	43	45.7	397	2 G83699	malate oxidoreduct
12	43	45.7	427	2 A43453	A-kinase anchor pr
13	43	45.7	461	2 G59093	hypothetical prote
14	43	45.7	542	2 S20466	hypothetical prote
15	43	45.7	1302	1 JG6009	surface-located me
16	42.5	45.2	2692	2 T23768	hypothetical prote
17	42	44.7	197	2 D91063	heat shock protein
18	42	44.7	213	2 H85907	heat shock protein
19	42	44.7	197	2 H57482	hypothetical prote
20	42	44.7	235	2 A56757	14-3-3 regulatory
21	42	44.7	347	2 T35518	probable secreted
22	42	44.7	353	2 G81120	I protein, probabl
23	42	44.7	420	2 D82709	colicin V secretio
24	42	44.7	649	2 T22254	hypothetical prote
25	42	44.7	722	2 H97217	uncharacterized co
26	42	44.7	769	2 T22256	hypothetical prote
27	42	44.7	1189	2 AH2154	WD-repeat protein
28	42	44.7	1668	2 A60272	IGA-specific metal
29	41	43.6	110	2 B40785	ORF2 protein - ric

30	41	43.6	110	2 S16667	hypothetical prote
31	41	43.6	197	2 S01240	heat shock protein
32	41	43.6	245	2 T26828	hypothetical prote
33	41	43.6	393	2 S02185	uroporphyrin-III C
34	41	43.6	399	2 E91220	uroporphyrinogen I
35	41	43.6	399	2 G86066	uroporphyrinogen I
36	41	43.6	414	2 T50010	hypothetical prote
37	41	43.6	731	2 S16595	gene CARS12 prote
38	41	43.6	1063	2 T18255	cytoskeleton assem
39	41	43.6	1171	2 T17454	diaphanous-related
40	41	43.6	1228	2 A57384	multimerin, endoth
41	41	43.6	1513	2 T44045	hypothetical prote
42	41	43.6	1520	2 T44231	hypothetical prote
43	40	42.6	67	2 A11987	hypothetical prote
44	40	42.6	184	2 H81850	probable membrane
45	40	42.6	184	2 F81090	hypothetical prote

ALIGNMENTS

RESULT 1

G95258

secreted 45 kd protein [imported] - Streptococcus pneumoniae (strain TIGR4)

C:Species: Streptococcus pneumoniae

C>Date: 03-Aug-2001 #sequence_revision 03-Aug-2001 #text_change 09-Jul-2004

C:Accession: G95258

R:Tetelin, H.; Nelson, K.E.; Paulsen, I.T.; Eisen, J.A.; Read, T.D.; Peterson, S.; Heirson, J.D.; Umayam, L.A.; White, O.; Salzberg, S.L.; Lewis, M.R.; Radune, D.; Holtzapple, nson, T.; Hickey, B.K.; Holt, I.E.

Science 293, 498-506, 2001

A:Authors: Loftus, B.J.; Yang, F.; Smith, H.O.; Venter, J.C.; Dougherty, B.A.; Morrison,

A>Title: Complete Genome Sequence of a virulent isolate of Streptococcus pneumoniae.

A:Reference number: A95000; MUID:21357209; PMID:11463916

A:Accession: G95258

A>Status: preliminary

A:Molecule type: DNA

A:Residues: 1-392 <KUR>

A:Cross-references: UNIPROT:Q97N55; UNIPROT:Q8DMY4; UNIPARC:UPI0000051B81; GB:AE005672;

A:Experimental source: strain TIGR4

C:Genetics:

A:Gene: SP2216

Query Match 78.7%; Score 74; DB 2; Length 392;
Best Local Similarity 75.0%; Pred. NO. 0.00026;
Matches 15; Conservative 4; Mismatches 1; Indels 0; Gaps 0;

QY 1 QIASQDSKINLTAQQQAAQ 20

DB 33 KIAAQDNKISNLTAQQQEAQ 52

RESULT 2

B98124

General stress protein GSP-781 [imported] - Streptococcus pneumoniae (strain R6)

C:Species: Streptococcus pneumoniae

C>Date: 22-Oct-2001 #sequence_revision 22-Oct-2001 #text_change 09-Jul-2004

C:Accession: B98124

R:Hoskins, J.A.; Alborn Jr., W.; Arnold, J.; Blaazczak, L.; Burgett, S.; DeHoff, B.S.; F e, R.; LeBlanc, D.J.; Lee, L.N.; Lefkowitz, E.J.; Lu, J.; Matsushima, P.; McAhren, S.; M y, P.; Sun, P.M.; Winkler, M.E.

J. Bacteriol. 183, 5709-5717, 2001

A:Authors: Yang, Y.; Young-Bellido, M.; Zhao, G.; Zook, C.; Baltz, R.H.; Jaskunas, S.R.;

A>Title: Genome of the Bacterium Streptococcus pneumoniae Strain R6.

A:Reference number: A97872; MUID:21429245; PMID:11544234

A:Accession: B98124

A>Status: preliminary

A:Molecule type: DNA

A:Residues: 1-392 <KUR>

A:Cross-references: UNIPROT:Q97N55; UNIPROT:Q8DMY4; UNIPARC:UPI0000051B81; GB:AE007317;

C:Genetics:

A:Gene: gsp-781

```
Query Match      78.7%; Score 74; DB 2; Length 392;
Best Local Similarity 75.0%; Pred. No. 0.00026;
Matches 15; Conservative 4; Mismatches 1; Indels 0; Gaps 0;

QY 1 QIASQDSKINNLTAAQQAAQ 20
   |||:|:|:|:|:|:|:|:|:|
Db 33 KIAAQDNKISNLTAAQQAEQ 52

RESULT 3
S05542
hypothetical protein, 54K - Enterococcus faecium
C:Species: Enterococcus faecium
C:Date: 07-Sep-1990 #sequence_revision 07-Sep-1990 #text_change 09-Jul-2004
A:Accession: S05542
R:Fuerst, P.; Moesch, H.U.; Solioz, M.
Nucleic Acids Res. 17, 6724, 1989
A:Title: A protein of unusual composition from Enterococcus faecium.
A:Reference number: S05542; MUID:89385998; PMID:2780297
A:Accession: S05542
A:Molecule type: DNA
A:Status: preliminary
A:Residues: 1-507 <FUE>
A:Cross-references: UNIPROT:P13692; UNIPARC:UPI000016F6PC; GB:X16421; EMBL:M26048; NID:3
A:Note: the authors translated the codon CGT for residues 221 and 223 as Lys

Query Match      50.0%; Score 47; DB 2; Length 507;
Best Local Similarity 55.0%; Pred. No. 8.7;
Matches 11; Conservative 2; Mismatches 7; Indels 0; Gaps 0;

QY 1 QIASQDSKINNLTAAQQAAQ 20
   |||:|:|:|:|:|:|:|:|:|
Db 24 QIQQQDKKIADLQNOQASQAQ 43

RESULT 4
D75477
hypothetical protein - Deinococcus radiodurans (strain R1)
C:Species: Deinococcus radiodurans
C:Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 09-Jul-2004
A:Accession: D75477
R:White, O.; Eisen, J.A.; Heidelberg, J.F.; Hickey, E.K.; Peterson, J.D.; Dodson, R.J.;
  M.; Shen, M.; Vanathavan, J.J.; Lam, P.; McDonald, L.; Utterback, T.; Zalewski, C.; Ma
  S.; Smith, H.O.; Venter, J.C.; Fraser, C.M.
Science 286, 1571-1577, 1999
A:Title: Genome sequence of the radioresistant bacterium Deinococcus radiodurans R1.
A:Reference number: A75250; MUID:20036896; PMID:10567266
A:Accession: D75477
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-238 <WHI>
A:Cross-references: UNIPROT:Q9RW98; UNIPARC:UPI00000C1802; GB:AE001932; GB:AE000513; NID
A:Experimental source: strain R1
C:Genetics:
A:Gene: DR0771
A:Map position: 1

Query Match      48.9%; Score 46; DB 2; Length 238;
Best Local Similarity 60.0%; Pred. No. 5.7;
Matches 9; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

QY 2 IASQDSKINNLTAAQ 16
   :|:|:|:|:|:|:|:|:|
Db 63 VAPQKSRFNDLTAQ 77

RESULT 5
AB2734
NADH dehydrogenase I chain M nuoM [imported] - Agrobacterium tumefaciens (strain C58, Du
C:Species: Agrobacterium tumefaciens
C:Date: 11-Jan-2002 #sequence_revision 11-Jan-2002 #text_change 09-Jul-2004
A:Accession: AB2734
R:Wood, D.W.; Setubal, J.C.; Kaul, R.; Monks, D.; Chen, L.; Wood, G.E.; Chen, Y.; Woo, I
erage, G.; Gillet, W.; Grant, C.; Guenther, D.; Kuttyavin, T.; Levy, R.; Li, M.; McClell
```

```
i Karp, P.; Romero, P.; Zhang, S.
Science 294, 2317-2323, 2001
A:Authors: Yoo, H.; Tao, Y.; Biddle, P.; Jung, M.; Kreepan, W.; Perry, M.; Gordon-Kamm,
ster, E.W.
A:Title: The Genome of the Natural Genetic Engineer Agrobacterium tumefaciens C58.
A:Reference number: AB2577; MUID:21608550; PMID:11743193
A:Accession: AB2734
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-503 <KUR>
A:Cross-references: UNIPROT:Q8UFW5; UNIPARC:UPI00000D1AD8; GB:AE008688; PIDN:AAL42288.1;
A:Experimental source: strain C58 (Dupont)
C:Genetics:
A:Gene: nuoM
A:Map position: circular chromosome
C:Superfamily: NADH dehydrogenase (ubiquinone) chain 4

Query Match      48.9%; Score 46; DB 2; Length 503;
Best Local Similarity 61.1%; Pred. No. 13;
Matches 11; Conservative 1; Mismatches 6; Indels 0; Gaps 0;

QY 3 ASQDSKINNLTAAQQAAQ 20
   |||:|:|:|:|:|:|:|:|:|
Db 479 ASVDLLVNNYTAALQAAQ 496

RESULT 6
D97515
NADH dehydrogenase I chain m (NADH-ubiquinone oxidoreductase chain m) AGR_C_2362 [import
C:Species: Agrobacterium tumefaciens
C:Date: 30-Sep-2001 #sequence_revision 30-Sep-2001 #text_change 09-Jul-2004
A:Accession: D97515
R:Goodner, B.; Hinkle, G.; Gattung, S.; Miller, N.; Blanchard, M.; Quorollo, B.; Goldman,
A.; Liu, F.; Wollam, C.; Allinger, M.; Doughty, D.; Scott, C.; Lappas, C.; Markelz, B.;
Science 294, 2323-2328, 2001
A:Title: Genome Sequence of the Plant Pathogen and Biotechnology Agent Agrobacterium tum
A:Reference number: A97359; MUID:21608551; PMID:11743194
A:Accession: D97515
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-503 <KUR>
A:Cross-references: UNIPROT:Q8UFW5; UNIPARC:UPI00000D1AD8; GB:AE007869; PIDN:AAK87077.1;
C:Genetics:
A:Gene: AGR_C_2362
A:Map position: circular chromosome
C:Superfamily: NADH dehydrogenase (ubiquinone) chain 4

Query Match      48.9%; Score 46; DB 2; Length 503;
Best Local Similarity 61.1%; Pred. No. 13;
Matches 11; Conservative 1; Mismatches 6; Indels 0; Gaps 0;

QY 3 ASQDSKINNLTAAQQAAQ 20
   |||:|:|:|:|:|:|:|:|:|
Db 479 ASVDLLVNNYTAALQAAQ 496

RESULT 7
E88940
protein C0584.9 [imported] - Caenorhabditis elegans
C:Species: Caenorhabditis elegans
C:Date: 10-May-2001 #sequence_revision 10-May-2001 #text_change 09-Jul-2004
A:Accession: E88940
R:anonymous, The C. elegans Sequencing Consortium.
Science 282, 2012-2018, 1998
A:Title: Genome sequence of the nematode C. elegans: a platform for investigating biolog
A:Reference number: A75000; MUID:99069613; PMID:9851916
A:Note: see websites genome.wustl.edu/gsc/C_elegans/ and www.sanger.ac.uk/Projects/C_eleg
A:Note: published errata appeared in Science 283, 35, 1999; Science 283, 2103, 1999; an
A:Accession: E88940
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-526 <STO>
A:Cross-references: UNIPROT:Q10663; UNIPARC:UPI0000176000; GB:chr_V; PIDN:AA871278.1; PI
```

C;Genetics:
A:Gene: C05E4.9
A:Map position: 5
C;Superfamily: isocitrate lyase

Query Match 47.9%; Score 45; DB 2; Length 526;
Best Local Similarity 58.8%; Pred. No. 19;
Matches 10; Conservative 3; Mismatches 4; Indels 0; Gaps 0;

Qy 2 IASQDSKINNLTAAQQQA 18
:||||:|:|:|:|
Db 435 VASQDEEILSLTAQNV 451

RESULT 8
AF1990
hypotheical protein all1475 [imported] - Nostoc sp. (strain PCC 7120)
C;Species: Nostoc sp. PCC 7120
A:Note: Nostoc sp. strain PCC 7120 is a synonym of Anabaena sp. strain PCC 7120
C;Date: 14-Dec-2001 #sequence_revision 14-Dec-2001 #text_change 09-Jul-2004
C;Accession: AF1990
R;Kaneko, T.; Nakamura, Y.; Wolk, C.P.; Kuritz, T.; Sasamoto, S.; Watanabe, A.; Iriiguchi, N.; Nakazaki, N.; Shimpo, S.; Sugimoto, M.; Takazawa, M.; Yamada, M.; Tabata, S.; Tabata, S.
DNA Res. 8, 205-213, 2001
A;Title: Complete Genomic Sequence of the Filamentous Nitrogen-fixing Cyanobacterium Anabaena
A;Reference number: AB1807; MUID:21595285; PMID:11759840
A;Accession: AF1990
A;Status: preliminary
A:Molecule type: DNA
A;Residues: 1-308 <KUR>
A;Cross-references: UNIPROT:Q9YWH9; UNIPARC:UPI000000CE0DC; GB:BA000019; PIDN:BA077841.1;
A;Experimental source: strain PCC 7120
C;Genetics:
A:Gene: all1475

Query Match 46.8%; Score 44; DB 2; Length 308;
Best Local Similarity 45.0%; Pred. No. 16;
Matches 9; Conservative 5; Mismatches 6; Indels 0; Gaps 0;

Qy 1 QIASQDSKINNLTAAQQQAQ 20
:||||:|:|:|:|
Db 88 EMASVKERIKNLTPQKEAQ 107

RESULT 9
S15194
sucrose porin scrY - Klebsiella pneumoniae
C;Species: Klebsiella pneumoniae
C;Date: 21-Nov-1993 #sequence_revision 26-May-1995 #text_change 09-Jul-2004
C;Accession: S15194
R;Schmid, K.; Ebner, R.; Jahreis, K.; Lengeler, J.W.; Titgemeyer, F.
Mol. Microbiol. 5, 941-950, 1991
A;Title: A sugar-specific porin, ScrY, is involved in sucrose uptake in enteric bacteria
A;Reference number: S15193; MUID:91312133; PMID:1649946
A;Accession: S15194
A;Status: preliminary; translation not shown
A:Molecule type: DNA
A;Residues: 1-505 <SCH>
A;Cross-references: UNIPROT:P27218; UNIPARC:UPI00001356AF; EMBL:X57401; NID:g43931; PIDN:
C;Genetics:
A:Gene: scrY

Query Match 46.8%; Score 44; DB 2; Length 505;
Best Local Similarity 50.0%; Pred. No. 27;
Matches 9; Conservative 2; Mismatches 7; Indels 0; Gaps 0;

Qy 3 ASQDSKINNLTAAQQQAQ 20
|||:|:|:|:|
Db 55 ASAEQKVQQLTQQQQQTQ 72

RESULT 10
T23107

A;Note: sequence extracted from NCBI backbone (NCBIN:111869, NCBIPI:111870)

Query Match 45.7%; Score 43; DB 2; Length 427;
Best Local Similarity 66.7%; Pred. No. 33;
Matches 8; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 1 QIASQDSKINNLT 12
::|||:|||||
DB 414 EWASDDNKINNLT 425

RESULT 13

G59093
hypothetical protein pX01-23 - Bacillus anthracis virulence plasmid pX01
C;Species: Bacillus anthracis
C;Date: 12-Nov-1999 #sequence_revision 12-Nov-1999 #text_change 09-Jul-2004
C;Accession: G59093
C;Okimaka, R.T.; Cloud, K.; Hampton, O.; Hoffmaster, A.R.; Hill, K.K.; Keim, P.; Koehler
J. Bacteriol. 181, 6509-6515, 1999
A;Title: Sequence and organization of pX01, the large Bacillus anthracis plasmid harbored
A;Reference number: AS9091; MUID:99445483; PMID:10515943
A;Accession: G59093
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-461 <OKI>
A;Cross-references: UNIPROT:Q9X224; UNIPARC:UPI00000B7358; GB:AF065404; NID:G4894216; PI
A;Experimental source: strain Sterne
A;Note: similar to hypothetical gene (609 aa) Clostridium difficile (X98606), 249/390 po
C;Genetics:
A;Gene: pX01-23
A;Genome: plasmid

Query Match 45.7%; Score 43; DB 2; Length 461;
Best Local Similarity 53.3%; Pred. No. 35;
Matches 8; Conservative 4; Mismatches 3; Indels 0; Gaps 0;

QY 2 IASQDSKINNLTAAQ 16
|:::|::|::|::|
DB 331 ISAEKSKVTLTKQQ 345

RESULT 14

S20466
hypothetical protein - fungus (Fusarium oxysporum)
C;Species: Fusarium oxysporum
C;Date: 06-Feb-1995 #sequence_revision 06-Feb-1995 #text_change 09-Jul-2004
C;Accession: S20466
R;Daboussi, M.J.; Langin, T.; Brygoo, Y.
Mol. Gen. Genet. 232, 12-16, 1992
A;Title: Pot1, a new family of fungal transposable elements.
A;Reference number: S20466; MUID:92204124; PMID:1313143
A;Accession: S20466
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-542 <DAB>
A;Cross-references: UNIPROT:Q00832; UNIPARC:UPI000006B3A2; EMBL:X64799; NID:G2722; PIDN:

Query Match 45.7%; Score 43; DB 2; Length 542;
Best Local Similarity 45.0%; Pred. No. 42;
Matches 9; Conservative 5; Mismatches 6; Indels 0; Gaps 0;

QY 1 QIASQDSKINNLTAAQ 20
::||:|::|::|::|
DB 438 ELAMKDREYERLRAQLAAQ 457

RESULT 15

JC6009
surface-located membrane protein lmp3 precursor - Mycoplasma hominis
C;Species: Mycoplasma hominis
C;Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 09-Jul-2004
C;Accession: JC6009
R;Ladefoged, S.A.; Jensen, L.T.; Brock, B.; Birkelund, S.; Christiansen, G.

J. Bacteriol. 178, 2775-2784, 1996
A;Title: Analysis of 0.5-kilobase-pair repeats in the Mycoplasma hominis lmp gene system
A;Reference number: JC6009; MUID:96213016; PMID:8631664
A;Accession: JC6009
A;Molecule type: DNA
A;Residues: 1-1302 <LAD>
A;Cross-references: UNIPROT:Q49547; UNIPARC:UPI00000B0501; EMBL:X95601; NID:gl197335; PI
C;Genetics:
A;Gene: lmp3
A;Genetic code: SGC3
A;Superfamily: surface-located membrane protein lmp3; tetratrico peptide repeat homology
C;Keywords: duplication; membrane protein
F;1-24/Domain: signal sequence #status predicted <SIG>
F;25-1302/Product: surface-located membrane protein lmp3 #status predicted <MAT>
F;957-992/Domain: tetratrico peptide repeat homology <TT1>
F;993-1026/Domain: tetratrico peptide repeat homology <TT2>
F;1089-1120/Domain: tetratrico peptide repeat homology <TT3>
F;1154-1190/Domain: tetratrico peptide repeat homology <TT4>

Query Match 45.7%; Score 43; DB 1; Length 1302;
Best Local Similarity 45.0%; Pred. No. 1.1e+02;
Matches 9; Conservative 4; Mismatches 7; Indels 0; Gaps 0;

QY 1 QIASQDSKINNLTAAQ 20
|:::|::|::|::|
DB 679 QSASQELNLTNNPTSMQSAKE 698

Search completed: January 25, 2006, 19:15:17
Job time : 9.1 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2006 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: January 25, 2006, 18:48:24 ; Search time 70.3 Seconds
(without alignments)
200.719 Million cell updates/sec

Title: US-10-797-821-6

Perfect score: 94

Sequence: 1 QIASQDSKINNLTAAQQAAQ 20

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 2166443 seqs, 705528306 residues

Total number of hits satisfying chosen parameters: 2166443

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : UniProt_05.80.*

1: uniprot_sprot.*

2: uniprot_trembl.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	ID	Description
1	94	100.0	Q938V0_STRMU	Q938V0 streptococc
2	94	100.0	Q938V3_STRMU	Q938V3 streptococc
3	94	100.0	Q9AG98_STRMU	Q9AG98 streptococc
4	94	100.0	Q8DMW3_STRMU	Q8DMW3 streptococc
5	94	100.0	Q938V1_STRMU	Q938V1 streptococc
6	94	100.0	Q938V2_STRMU	Q938V2 streptococc
7	81	86.2	Q9ZAS7_STRMU	Q9ZAS7 streptococc
8	74	78.7	Q8DMY4_STRMR6	Q8DMY4 streptococc
9	74	78.7	Q97N55_STRPN	Q97N55 streptococc
10	62	66.0	Q5XEL1_STRP6	Q5XEL1 streptococc
11	62	66.0	Q9ALZ8_STRPY	Q9ALZ8 streptococc
12	62	66.0	Q7CNQ7_STRP8	Q7CNQ7 streptococc
13	62	66.0	Q8P318_STRP3	Q8P318 streptococc
14	55	58.5	Q5M6K4_STRT2	Q5M6K4 streptococc
15	55	58.5	Q5M212_STRT1	Q5M212 streptococc
16	55	58.5	Q5SEG2_DICD1	Q5SEG2 dictyosteli
17	54	57.4	Q9AKA4_STRAG	Q9AKA4 streptococc
18	54	57.4	Q8E2H1_STRAS	Q8E2H1 streptococc
19	54	57.4	Q8E7X9_STRAS3	Q8E7X9 streptococc
20	51.5	54.8	Q4P224_USTMA	Q4P224 ustilago ma
21	49	52.1	Q70P98_9DEL1	Q70P98 melittangiu
22	48	51.1	Q810C2_DROME	Q810C2 drosophila
23	48	51.1	Q9VN98_DROME	Q9VN98 drosophila
24	48	51.1	Q8SX21_DROME	Q8SX21 drosophila
25	48	51.1	Q5ZIN3_NOCFJA	Q5ZIN3 nocardia fa
26	48	51.1	Q74LW7_LACJO	Q74LW7 lactobacill
27	48	51.1	Q7ZAJ7_SHEON	Q7ZAJ7 shewanella
28	48	51.1	Q7Z8J7_USTMA	Q7Z8J7 ustilago ma
29	48	51.1	Q4P5Q8_USTMA	Q4P5Q8 ustilago ma
30	47	50.0	Q6RKH4_GIBBEZ	Q6RKH4 gibberella
31	47	50.0	Q9AZ34_BPHK6	Q9AZ34 bacterioph

32	47	50.0	482	2	Q93LK4_ENTFA	Q93LK4 enterococcu
33	47	50.0	516	1	P54_ENTFC	P13692 enterococcu
34	47	50.0	576	2	Q9KJ33_ENTHR	Q9KJ33 enterococcu
35	47	50.0	618	2	Q9VZB5_DROME	Q9VZB5 drosophila
36	47	50.0	797	2	Q514B8_ENTHI	Q514B8 entamoeba h
37	47	50.0	2298	2	Q5T2I9_BRARE	Q5T2I9 brachydanio
38	47	50.0	2612	2	Q413K0_GIBZE	Q413K0 gibberella
39	46	48.9	238	2	Q9RW98_DEIRA	Q9RW98 deinococcus
40	46	48.9	244	2	Q5XGC8_XENTR	Q5XGC8 xenopus tro
41	46	48.9	503	2	Q8UFW5_AGR15	Q8UFW5 agrobacteri
42	45	47.9	110	2	Q8V6W1_9VIRU	Q8V6W1 rice tungro
43	45	47.9	110	2	Q9WQX7_9VIRU	Q9WQX7 rice tungro
44	45	47.9	110	2	Q9WQV3_9VIRU	Q9WQV3 rice tungro
45	45	47.9	110	2	Q9YRV5_9VIRU	Q9YRV5 rice tungro

ALIGNMENTS

RESULT 1

Q938V0_STRMU	PRELIMINARY;	PRT;	431 AA.
ID Q938V0;			
AC Q938V0;			
DT 01-DEC-2001 (Tremblrel. 19, Created)			
DT 01-DEC-2001 (Tremblrel. 19, Last sequence update)			
DT 01-MAR-2004 (Tremblrel. 26, Last annotation update)			
DE Glucan-binding protein B.			
OS Streptococcus mutans.			
OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;			
OC Streptococcus.			
OX NCBI_TaxID=1309;			
RN [1]			
RP NUCLEOTIDE SEQUENCE.			
RC STRAIN=SSM3;			
RA Jin S., Duncan M.J., Taubman M.A., Smith D.J.;			
RT "Cloning of the gbpB gene from Streptococcus mutans.";			
RL J. Dent. Res. 79:224-224(2000).			
RN [2]			
RP NUCLEOTIDE SEQUENCE.			
RC STRAIN=SSM3;			
RX MEDLINE=21481971; PubMed=11598068;			
RX DOI=10.1128/IAI.69.11.6931-6941.2001;			
RA Mattos-Graner R.O., Jin S., King W.F., Chen T., Smith D.J.,			
Duncan M.J.;			
RT "Cloning of the Streptococcus mutans gene encoding glucan binding			
protein B and analysis of genetic diversity and protein production in			
clinical isolates.";			
RL Infect. Immun. 69:6931-6941(2001).			
DR EMBL; AY046414; AAK94504.1; -; Genomic_DNA.			
DR InterPro; IPR007921; CHAP.			
DR InterPro; IPR009148; Siba.			
DR Pfam; PF05257; CHAP; 1.			
DR PRINTS; PR01852; SIRAPROTEIN.			
DR PROSITE; PS50911; CHAP; 1.			
SQ SEQUENCE 431 AA; 44650 MW; 05D38D8B8C4609F CRC64;			

Query Match 100.0%; Score 94; DB 2; Length 431;
Best Local Similarity 100.0%; Pred. No. 1.9e-06;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY	1	QIASQDSKINNLTAAQQAAQ 20
DB	33	QIASQDSKINNLTAAQQAAQ 52

RESULT 2

Q938V3_STRMU	PRELIMINARY;	PRT;	431 AA.
ID Q938V3;			
AC Q938V3;			
DT 01-DEC-2001 (Tremblrel. 19, Created)			
DT 01-DEC-2001 (Tremblrel. 19, Last sequence update)			
DT 01-MAR-2004 (Tremblrel. 26, Last annotation update)			
DE Glucan-binding protein B.			

OS Streptococcus mutans.
 OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
 OC Streptococcus.
 OX NCBI_TaxID=1309;

[1]

RP NUCLEOTIDE SEQUENCE.

RC STRAIN=GS-5;

RA Jia J.S., Chang L.Y., Shun C.T., Chang Y.Y., Chen J.Y.;

RT "Cloning of the gbpB gene from Streptococcus mutans.";

RL J. Dent. Res. 79:224-224(2000).

[2]

RP NUCLEOTIDE SEQUENCE.

RC STRAIN=GS-5;

RA Jia J.S., Chang L.Y., Shun C.T., Chang Y.Y., Chen J.Y.;

RT "Cloning of the gbpB gene from Streptococcus mutans.";

RL J. Dent. Res. 79:224-224(2000).

RA Jia J.S., Chang L.Y., Shun C.T., Chang Y.Y., Chen J.Y.;

RT "Cloning of the Streptococcus mutans gene encoding glucan binding

protein B and analysis of genetic diversity and protein production in

clinical isolates.";

RL Infect. Immun. 69:6931-6941(2001).

DR EMBL; AY046410; AAK94500.1; -; Genomic_DNA.

DR InterPro; IPR007921; CHAP.

DR Pfam; PF05257; CHAP; 1.

DR PRINTS; PR01852; SIBAPROTEIN.

DR PROSITE; PS50911; CHAP; 1.

SQ SEQUENCE 431 AA; 44620 MW; 464FE3B563FB7E51 CRC64;

Query Match 100.0%; Score 94; DB 2; Length 431;

Best Local Similarity 100.0%; Pred. No. 1.9e-06;

Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 QIASQDSKINNLTAAQQAAQ 20

|||||

Db 33 QIASQDSKINNLTAAQQAAQ 52

RESULT 3

Q9AG98 STRMU

ID Q9AG98 STRMU PRELIMINARY; PRT; 431 AA.

AC Q9AG98;

DT 01-JUN-2001 (TrEMBLrel. 17, Created)

DT 01-FEB-2005 (TrEMBLrel. 17, Last sequence update)

DT 01-FEB-2005 (TrEMBLrel. 29, Last annotation update)

DE Immunodominant glycoprotein IDG-60 (Glucan-binding protein B).

GN Name=sagA;

OS Streptococcus mutans.

OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;

OC Streptococcus.

OX NCBI_TaxID=1309;

RN [1]

RP NUCLEOTIDE SEQUENCE.

RC STRAIN=GS-5;

RA Jia J.S., Chang L.Y., Shun C.T., Chang Y.Y., Chen J.Y.;

RT "Identification of stress-responsive genes in Streptococcus mutans by

differential display reverse transcription-PCR.";

RL Infect. Immun. 69:2493-2501(2001).

RN [2]

RP NUCLEOTIDE SEQUENCE.

RC STRAIN=GS-5;

RA Jia J.S., Chang L.Y., Shun C.T., Chang Y.Y., Chen J.Y.;

RT "A 60-kilodalton immunodominant glycoprotein is essential for cell

wall integrity and the maintenance of cell shape in Streptococcus

mutans.";

RL Infect. Immun. 69:6987-6998(2001).

RN [3]

RP NUCLEOTIDE SEQUENCE.

RC STRAIN=GS-5;

RA Jia J.S., Chang L.Y., Shun C.T., Chang Y.Y., Chen J.Y.;

RT "Cloning of the Streptococcus mutans gene encoding glucan binding

protein B and analysis of genetic diversity and protein production in

clinical isolates.";

RL Infect. Immun. 69:6931-6941(2001).

DR EMBL; AY046410; AAK94500.1; -; Genomic_DNA.

DR InterPro; IPR007921; CHAP.

DR Pfam; PF05257; CHAP; 1.

DR PRINTS; PR01852; SIBAPROTEIN.

DR PROSITE; PS50911; CHAP; 1.

SQ SEQUENCE 431 AA; 44620 MW; 464FE3B563FB7E51 CRC64;

RA Chia J.-S., Chang L.-Y., Lee Y.-Y., Chen J.-Y.;

RL Submitted (NOV-2001) to the EMBL/GenBank/DBJ databases.

[4]

RP NUCLEOTIDE SEQUENCE.

RC STRAIN=3VF4;

RA Jin S., Duncan M.J., Taubman M.A., Smith D.J.;

RT "Cloning of the gbpB gene from Streptococcus mutans.";

RL J. Dent. Res. 79:224-224(2000).

[5]

RP NUCLEOTIDE SEQUENCE.

RC STRAIN=3VF4;

RA Jin S., Duncan M.J., Taubman M.A., Smith D.J.;

RT "Cloning of the gbpB gene from Streptococcus mutans.";

RL J. Dent. Res. 79:224-224(2000).

RA Jin S., Duncan M.J., Taubman M.A., Smith D.J.;

RT "Cloning of the Streptococcus mutans gene encoding glucan binding

protein B and analysis of genetic diversity and protein production in

clinical isolates.";

RL Infect. Immun. 69:6931-6941(2001).

DR EMBL; AF338445; AAK08104.1; -; Genomic_DNA.

DR EMBL; AY046411; AAK94501.1; -; Genomic_DNA.

DR InterPro; IPR007921; CHAP.

DR Pfam; PF05257; CHAP; 1.

DR PRINTS; PR01852; SIBAPROTEIN.

DR PROSITE; PS50911; CHAP; 1.

SQ SEQUENCE 431 AA; 44592 MW; 3EBE21FC5E47232E CRC64;

Query Match 100.0%; Score 94; DB 2; Length 431;

Best Local Similarity 100.0%; Pred. No. 1.9e-06;

Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 QIASQDSKINNLTAAQQAAQ 20

|||||

Db 33 QIASQDSKINNLTAAQQAAQ 52

RESULT 4

Q8DWM3 STRMU

ID Q8DWM3 STRMU PRELIMINARY; PRT; 431 AA.

AC Q8DWM3;

DT 01-MAR-2003 (TrEMBLrel. 23, Created)

DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)

DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)

DE Putative secreted antigen GbpB/SagA; putative peptidoglycan

hydrolyase.

GN Name=gbpB; OrderedLocName=SMU.22;

OS Streptococcus mutans.

OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;

OC Streptococcus.

OX NCBI_TaxID=1309;

RN [1]

RP NUCLEOTIDE SEQUENCE.

RC STRAIN=UA159 / ATCC 700610 / Serotype c;

RA Jia J.S., Chang L.Y., Shun C.T., Chang Y.Y., Chen J.Y.;

RT "Genome sequence of Streptococcus mutans UA159, a cariogenic dental

pathogen.";

RL Proc. Natl. Acad. Sci. U.S.A. 99:14434-14439(2002).

DR EMBL; AE014855; AAN57811.1; -; Genomic_DNA.

DR InterPro; IPR007921; CHAP.

DR Pfam; PF05257; CHAP; 1.

DR PRINTS; PR01852; SIBAPROTEIN.

DR PROSITE; PS50911; CHAP; 1.

KW Complete proteome.

SQ SEQUENCE 431 AA; 44620 MW; 2D1CA685248CCD3E CRC64;

Query Match 100.0%; Score 94; DB 2; Length 431;

Best Local Similarity 100.0%; Pred. No. 1.9e-06;

Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 QIASQDSKINNLTAQQAQ 20
DB 33 QIASQDSKINNLTAQQAQ 52

RESULT 5

Q938V1 STRMU
ID Q938V1 STRMU PRELIMINARY; PRT; 432 AA.
AC Q938V1;
DT 01-DEC-2001 (Tremblrel. 19, Created)
DT 01-DEC-2001 (Tremblrel. 19, Last sequence update)
DT 01-MAR-2004 (Tremblrel. 26, Last annotation update)
DE Glucan-binding protein B.
OS Streptococcus mutans.
OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
OC Streptococcus.
OX NCBI_TaxID=1309;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=3SN1;
RA Jin S., Duncan M.J., Taubman M.A., Smith D.J.;
RT "Cloning of the gbpB gene from Streptococcus mutans.";
RL J. Dent. Res. 79:224-224(2000).
RN [2]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=3SN1;
RX MEDLINE=21481971; PubMed=11598068;
RX DOI=10.1128/IAI.69.11.6931-6941.2001;
RA Mattos-Graner R.O., Jin S., King W.F., Chen T., Smith D.J.,
RA Duncan M.J.;
RT "Cloning of the Streptococcus mutans gene encoding glucan binding
RT protein B and analysis of genetic diversity and protein production in
RT clinical isolates.";
RL Infect. Immun. 69:6931-6941(2001).
DR EMBL; AY046413; AAK94503.1; -; Genomic_DNA.
DR InterPro; IPR007921; CHAP.
DR Pfam; PF05257; CHAP; 1.
DR PRINTS; PR01852; SIBAPROTEIN.
DR PROSITE; PS50911; CHAP; 1.
SQ SEQUENCE 432 AA; 44652 MW; 3F88ECB9A1F3BE4F CRC64;

Query Match 100.0%; Score 94; DB 2; Length 432;
Best Local Similarity 100.0%; Pred. No. 1.9e-06;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 QIASQDSKINNLTAQQAQ 20
DB 33 QIASQDSKINNLTAQQAQ 52

RESULT 6

Q938V2 STRMU
ID Q938V2 STRMU PRELIMINARY; PRT; 432 AA.
AC Q938V2;
DT 01-DEC-2001 (Tremblrel. 19, Created)
DT 01-DEC-2001 (Tremblrel. 19, Last sequence update)
DT 01-MAR-2004 (Tremblrel. 26, Last annotation update)
DE Glucan-binding protein B.
OS Streptococcus mutans.
OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
OC Streptococcus.
OX NCBI_TaxID=1309;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=15JP2;
RA Jin S., Duncan M.J., Taubman M.A., Smith D.J.;
RT "Cloning of the gbpB gene from Streptococcus mutans.";
RL J. Dent. Res. 79:224-224(2000).
RN [2]
RP NUCLEOTIDE SEQUENCE.

RC STRAIN=15JP2;
RX MEDLINE=21481971; PubMed=11598068;
RX DOI=10.1128/IAI.69.11.6931-6941.2001;
RA Mattos-Graner R.O., Jin S., King W.F., Chen T., Smith D.J.,
RA Duncan M.J.;
RT "Cloning of the Streptococcus mutans gene encoding glucan binding
RT protein B and analysis of genetic diversity and protein production in
RT clinical isolates.";
RL Infect. Immun. 69:6931-6941(2001).
DR EMBL; AY046412; AAK94502.1; -; Genomic_DNA.
DR InterPro; IPR007921; CHAP.
DR Pfam; PF05257; CHAP; 1.
DR PRINTS; PR01852; SIBAPROTEIN.
DR PROSITE; PS50911; CHAP; 1.
SQ SEQUENCE 432 AA; 44648 MW; E769B2504AE50E9 CRC64;
Query Match 100.0%; Score 94; DB 2; Length 432;
Best Local Similarity 100.0%; Pred. No. 1.9e-06;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 QIASQDSKINNLTAQQAQ 20
DB 33 QIASQDSKINNLTAQQAQ 52

RESULT 7

Q9ZAS7 STRMU
ID Q9ZAS7 STRMU PRELIMINARY; PRT; 211 AA.
AC Q9ZAS7;
DT 01-MAY-1999 (Tremblrel. 10, Created)
DT 01-MAY-1999 (Tremblrel. 10, Last sequence update)
DT 01-OCT-2002 (Tremblrel. 22, Last annotation update)
DE Putative secreted protein (Fragment).
OS Streptococcus mutans.
OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
OC Streptococcus.
OX NCBI_TaxID=1309;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=GS-5 Kuramitsu;
RA Peruzzi F., Piggot P.J., Daneo-Moore L.;
RL Submitted (NOV-1996) to the EMBL/GenBank/DBJ databases.
DR EMBL; U78607; AAD00288.1; -; Genomic_DNA.
FT NON_TER 211 211
SQ SEQUENCE 211 AA; 22803 MW; 4ACE331159CFAFC6 CRC64;

Query Match 86.2%; Score 81; DB 2; Length 211;
Best Local Similarity 100.0%; Pred. No. 0.00011;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 QIASQDSKINNLTAQQAQ 17
DB 33 QIASQDSKINNLTAQQAQ 49

RESULT 8

Q8DMV4 STRR6
ID Q8DMV4 STRR6 PRELIMINARY; PRT; 392 AA.
AC Q8DMV4;
DT 01-MAR-2003 (Tremblrel. 23, Created)
DT 01-MAR-2003 (Tremblrel. 23, Last sequence update)
DT 01-MAR-2004 (Tremblrel. 26, Last annotation update)
DE General stress protein GSP-781.
GN Name=gsp-781; OrderedLocusNames=spr2021;
OS Streptococcus pneumoniae (strain ATCC BAA-255 / R6).
OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
OC Streptococcus.
OX NCBI_TaxID=171101;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC MEDLINE=21429245; PubMed=11544234;
RX DOI=10.1128/JB.183.19.5709-5717.2001;

RA Hoskins J., Alborn W.E. Jr., Arnold J., Blaszcak L.C., Burgett S.,
 RA DeHoff B.S., Estrem S.T., Fritz L., Fu D.-J., Fuller W., Geringer C.,
 RA Gilmore R., Glass J.S., Khoja H., Kraft A.R., Lagace R.E.,
 RA LeBlanc D.J., Lee L.N., Lefkowitz E.J., Lu J., Matsushima P.,
 RA McAhren S.M., McHenry M., McLeaster K., Mundy C.W., Nicase T.I.,
 RA Norris F.H., O'Gara M., Peery R.B., Robertson G.T., Rocky P.,
 RA Sun P.-M., Winkler M.E., Yang Y., Young-Bellido M., Zhao G.,
 RA Zook C.A., Baltz R.H., Jaskunas S.R., Rostek P.R. Jr., Skatrud P.L.,
 RA Glass J.I.;
 RT "Genome of the bacterium *Streptococcus pneumoniae* strain R6.";
 RL J. Bacteriol. 183:5709-5717(2001).
 DR EMBL; AE008566; AAL00823.1; -; Genomic_DNA.
 DR PIR; B98124; B98124.
 DR PIR; G95258; G95258.
 DR InterPro; IPR007921; CHAP.
 DR InterPro; IPR009148; Siba.
 DR Pfam; PF05257; CHAP; 1.
 DR PRINTS; PR01852; SIBAPROTEIN.
 DR PROSITE; PS50911; CHAP; 1.
 DR Complete proteome.
 KW SEQUENCE 392 AA; 41697 MW; B5A5A860FAEA77DD CRC64;

Query Match 78.7%; Score 74; DB 2; Length 392;
 Best Local Similarity 75.0%; Pred. No. 0.0028;
 Matches 15; Conservative 4; Mismatches 1; Indels 0; Gaps 0;

QY 1 QIASQDSKINNLTAAQQAQ 20
 :||:||||:|||||
 Db 33 KIAAQDNKISNLTAAQQAQ 52

RESULT 9

Q97N55 STPNP
 ID Q97N55 STPNP PRELIMINARY; PRT; 392 AA.
 AC Q97N55;
 DT 01-OCT-2001 (TrEMBLrel. 18, Created)
 DT 01-OCT-2001 (TrEMBLrel. 18, Last sequence update)
 DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
 DE Secreted 45 kd protein.
 GN OrderedLocustNames=SP2216;
 OS *Streptococcus pneumoniae*.
 OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
 OC Streptococcus.
 OX NCBI_TaxID=1313;
 RN [1]
 RP NUCLEOTIDE SEQUENCE.

RC STRAIN=ATCC BAA-334 / TIGR4;
 RX MEDLINE=21357209; PubMed=11463916; DOI=10.1126/science.1061217;
 RA Tettelin H., Nelson K.E., Paulsen I.T., Eisen J.A., Read T.D.,
 RA Peterson S.N., Heidelberg J.F., DeBoy R.T., Haft D.H., Dodson R.J.,
 RA Durkin A.S., Gwinn M.B., Kolonay J.F., Nelson W.C., Peterson J.D.,
 RA Umayam L.A., White O., Salzberg S.L., Lewis M.R., Radune D.,
 RA Holtzapple E.K., Khouri H.M., Wolf A.M., Utterback T.R., Hansen C.L.,
 RA McDonald L.A., Feldblyum T.V., Angiuoli S.V., Dickinson T.,
 RA Hickey E.K., Holt I.E., Loftus B.J., Yang F., Smith H.O., Venter J.C.,
 RA Dougherty B.A., Morrison D.A., Hollingshead S.K., Fraser C.M.;
 RT "Complete genome sequence of a virulent isolate of *Streptococcus pneumoniae*.";
 RL Science 293:498-506(2001).
 DR EMBL; AE007509; AAK76264.1; -; Genomic_DNA.
 DR PIR; B98124; B98124.
 DR PIR; G95258; G95258.
 DR TIGR; SP2216; -.
 DR InterPro; IPR007921; CHAP.
 DR InterPro; IPR009148; Siba.
 DR Pfam; PF05257; CHAP; 1.
 DR PRINTS; PR01852; SIBAPROTEIN.
 DR PROSITE; PS50911; CHAP; 1.
 DR Complete proteome.
 KW SEQUENCE 392 AA; 41697 MW; B5A5A860FAEA77DD CRC64;

Query Match 78.7%; Score 74; DB 2; Length 392;
 Best Local Similarity 75.0%; Pred. No. 0.0028;

Matches 15; Conservative 4; Mismatches 1; Indels 0; Gaps 0;
 QY 1 QIASQDSKINNLTAAQQAQ 20
 :||:||||:|||||
 Db 33 KIAAQDNKISNLTAAQQAQ 52

RESULT 10

Q9XEL1 STRP6
 ID Q5XEL1 STRP6 PRELIMINARY; PRT; 398 AA.
 AC Q5XEL1;
 DT 25-OCT-2004 (TrEMBLrel. 28, Created)
 DT 25-OCT-2004 (TrEMBLrel. 28, Last sequence update)
 DT 25-OCT-2004 (TrEMBLrel. 28, Last annotation update)
 DE Secreted protein.
 GN OrderedLocustNames=M6_Spy0017;
 OS *Streptococcus pyogenes* (serotype M6).
 OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
 OC Streptococcus.
 OX NCBI_TaxID=301450;
 RN [1]
 RP NUCLEOTIDE SEQUENCE.

RC STRAIN=MGAS10394;
 RX PubMed=15272401; DOI=10.1086/422697;
 RA Banks D.J., Porcella S.F., Barbican K.D., Beres S.B., Phillips L.E.,
 RA Voyich J.M., DeLeo F.R., Martin J.M., Somerville G.A., Musser J.M.;
 RT "Progress toward characterization of the group A *Streptococcus* metagenome: complete genome sequence of a macrolide-resistant serotype M6 strain.";
 RT J. Infect. Dis. 190:727-738(2004).
 DR EMBL; CP000003; AAT86152.1; -; Genomic_DNA.
 DR InterPro; IPR007921; CHAP.
 DR InterPro; IPR009148; Siba.
 DR Pfam; PF05257; CHAP; 1.
 DR PRINTS; PR01852; SIBAPROTEIN.
 DR PROSITE; PS50911; CHAP; 1.
 KW Complete proteome.
 SQ SEQUENCE 398 AA; 42028 MW; 5DCEDA78CB863B60 CRC64;

Query Match 66.0%; Score 62; DB 2; Length 398;
 Best Local Similarity 65.0%; Pred. No. 0.24;
 Matches 13; Conservative 4; Mismatches 3; Indels 0; Gaps 0;

QY 1 QIASQDSKINNLTAAQQAQ 20
 :||:||||:|||||
 Db 30 KIAAQDSKISNLTAAQQAQ 49

RESULT 11

Q9A1Z8 STRPY
 ID Q9A1Z8 STRPY PRELIMINARY; PRT; 398 AA.
 AC Q9A1Z8; Q7BH59;
 DT 01-JUN-2001 (TrEMBLrel. 17, Created)
 DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
 DT 10-MAY-2005 (TrEMBLrel. 30, Last annotation update)
 DE Secreted protein Siba precursor.
 GN OrderedLocustNames=Spy0019;
 OS *Streptococcus pyogenes*.
 OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
 OC Streptococcus.
 OX NCBI_TaxID=1314;
 RN [1]
 RP NUCLEOTIDE SEQUENCE.

RC STRAIN=SF370 / ATCC 700294 / Serotype M1;
 RX MEDLINE=21192684; PubMed=11296296; DOI=10.1073/pnas.071559398;
 RA Ferretti J.J., McShan W.M., Ajdic D.J., Savic D.J., Savic G., Lyon K.,
 RA Primeaux C., Sezate S., Suvorov A.N., Kenton S., Lai H.S., Lin S.P.,
 RA Qian Y., Jia H.G., Nejar F.Z., Ren Q., Zhu H., Song L., White J.,
 RA Yuan X., Clifford S.W., Roe B.A., McLaughlin R.E.;
 RT "Complete genome sequence of an M1 strain of *Streptococcus pyogenes*.";
 RL Proc. Natl. Acad. Sci. U.S.A. 98:4658-4663(2001).
 RN [2]
 RP NUCLEOTIDE SEQUENCE.

RA Fagan P.K., Reinscheid D., Gottschalk B., Chhatwal G.S.;
RT "Identification and characterization of a novel secreted protein from
group A streptococcus";
RL Submitted (NOV-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AR006474; AAK3158.1; -; Genomic DNA.
DR EMBL; AF319999; AAL73135.1; -; Genomic DNA.
DR InterPro: IPR007921; CHAP.
DR InterPro: IPR009148; SIBA.
DR Pfam: PF05257; CHAP; 1.
DR PRINTS; PR01852; SIBAPROTEIN.
DR PROSITE; PS0911; CHAP; 1.
KW Complete proteome; Signal.
FT SIGNAL 1 23 Potential
SQ SEQUENCE 398 AA; 41899 MW; 28A9B3F7195E969B CRC64;

Query Match 66.0%; Score 62; DB 2; Length 398;
Best Local Similarity 65.0%; Pred. No. 0.24;
Matches 13; Conservative 4; Mismatches 3; Indels 0; Gaps 0;

QY 1 QIASQDSKINNLTAAQQAQ 20
DB 30 KIAQDSIISNLTTEQKAQ 49

RESULT 12
QYCNQ7 STRP8 PRELIMINARY; PRT; 398 AA.
AC Q7CNQ7;
DT 05-JUL-2004 (TrEMBLrel. 27, Created)
DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE Putative secreted protein.
GN OrderedLocusNames=spym18_0020;
OS Streptococcus pyogenes (serotype M18).
OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
OC Streptococcus.
OX NCBI_TaxID=301451;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=MGAS8232 / Serotype M18;
RX MEDLINE=21927593; PubMed=11917108; DOI=10.1073/pnas.062526099;
RA Smoot J.C., Barbican K.D., Van Gompel J.J., Smoot L.M., Chaussee M.S.,
RA Sylva G.L., Sturdevant D.E., Ricklefs S.M., Porcella S.F.,
RA Parkins L.D., Beres S.B., Campbell D.S., Smith T.M., Zhang Q.,
RA Kapur V., Daly J.A., Veasy L.G., Musser J.M.;
RT "Genome sequence and comparative microarray analysis of serotype M18
group A Streptococcus strains associated with acute rheumatic fever
outbreaks";
RL Proc. Natl. Acad. Sci. U.S.A. 99:4668-4673 (2002).
DR EMBL; AE009955; AAL96849.1; -; Genomic DNA.
DR InterPro: IPR007921; CHAP.
DR InterPro: IPR009148; SIBA.
DR Pfam: PF05257; CHAP; 1.
DR PRINTS; PR01852; SIBAPROTEIN.
DR PROSITE; PS0911; CHAP; 1.
KW Complete proteome.
SQ SEQUENCE 398 AA; 42028 MW; 5DCEDA78CB863B60 CRC64;

Query Match 66.0%; Score 62; DB 2; Length 398;
Best Local Similarity 65.0%; Pred. No. 0.24;
Matches 13; Conservative 4; Mismatches 3; Indels 0; Gaps 0;

QY 1 QIASQDSKINNLTAAQQAQ 20
DB 30 KIAQDSIISNLTTEQKAQ 49

RESULT 13
QYCNQ7 STRP3 PRELIMINARY; PRT; 398 AA.
AC Q8P318;
DT 01-OCT-2002 (TrEMBLrel. 22, Created)
DT 01-OCT-2002 (TrEMBLrel. 22, Last sequence update)

DT 13-SEP-2005 (TrEMBLrel. 31, Last annotation update)
DE Putative secreted protein.
GN OrderedLocusNames=SPA0015, SpyM3_0014;
OS Streptococcus pyogenes (serotype M3).
OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
OC Streptococcus.
OX NCBI_TaxID=301448;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=SSI-1 / Serotype M3;
RX MEDLINE=22683278; PubMed=12799345; DOI=10.1101/gr.1096703;
RA Nakagawa I., Kurokawa K., Yamashita A., Nakata M., Tomiyasu Y.,
RA Okahashi N., Kawabata S., Yamazaki K., Shiba T., Yasunaga T.,
RA Hayashi H., Hattori M., Hamada S.;
RT "Genome sequence of an M3 strain of Streptococcus pyogenes reveals a
large-scale genomic rearrangement in invasive strains and new insights
into phage evolution";
RL Genome Res. 13:1042-1055 (2003).
RN [2]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=MGAS315 / Serotype M3;
RX MEDLINE=22133808; PubMed=12122206; DOI=10.1073/pnas.152298499;
RA Beres S.B., Sylva G.L., Barbican K.D., Lei B., Hoff J.S.,
RA Mammarella N.D., Liu M.-Y., Smoot J.C., Porcella S.F., Parkins L.D.,
RA Campbell D.S., Smith T.M., McCormick J.K., Leung D.Y.M.,
RA Schlievert P.M., Musser J.M.;
RT "Genome sequence of a serotype M3 strain of group A Streptococcus:
phage-encoded toxins, the high-virulence phenotype, and clone
emergence";
RL Proc. Natl. Acad. Sci. U.S.A. 99:10078-10083 (2002).
DR EMBL; BA000034; BAC63110.1; -; Genomic DNA.
DR EMBL; AE014136; AAM78621.1; -; Genomic DNA.
DR InterPro: IPR007921; CHAP.
DR InterPro: IPR009148; SIBA.
DR Pfam: PF05257; CHAP; 1.
DR PRINTS; PR01852; SIBAPROTEIN.
DR PROSITE; PS0911; CHAP; 1.
KW Complete proteome.
SQ SEQUENCE 398 AA; 42028 MW; 5DCEDA78CB863B60 CRC64;

Query Match 66.0%; Score 62; DB 2; Length 398;
Best Local Similarity 65.0%; Pred. No. 0.24;
Matches 13; Conservative 4; Mismatches 3; Indels 0; Gaps 0;

QY 1 QIASQDSKINNLTAAQQAQ 20
DB 30 KIAQDSIISNLTTEQKAQ 49

RESULT 14
QYCNQ7 STRT2 PRELIMINARY; PRT; 474 AA.
AC Q5M6K4;
DT 01-FEB-2005 (TrEMBLrel. 29, Created)
DT 01-FEB-2005 (TrEMBLrel. 29, Last sequence update)
DT 13-SEP-2005 (TrEMBLrel. 31, Last annotation update)
DE Glucan binding protein (PcsB).
GN Name=PcsB; OrderedLocusNames=stu0022;
OS Streptococcus thermophilus (strain ATCC BAA-250 / LMG 18311).
OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
OC Streptococcus.
OX NCBI_TaxID=264199;
RN [1]
RP NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
RX PubMed=15543133; DOI=10.1038/nbt1034;
RA Bolotin A., Quinquis B., Renault P., Sorokin A., Ehrlich S.D.,
RA Kulakauskas S., Lapidus A., Goitsman E., Mazur M., Pusch G.D.,
RA Fonstein M., Overbeek R., Kyprides N., Purnelle B., Prozzi D.,
RA Ngai K., Masuy D., Hancy F., Burtan S., Boutry M., Delcour J.,
RA Goffeau A., Hols P.;
RT "Complete sequence and comparative genome analysis of the dairy
bacterium Streptococcus thermophilus";
RL Nat. Biotechnol. 22:1554-1558 (2004).

```

RN NUCLEOTIDE SEQUENCE.
RP STRAIN=LMG 18311;
RA Borges F., Layec S., Thibessard A., Fernandez A., Gintz B., Hols P.,
RA Decaris B., Leblond-Bourget N.;
RT "cse, a Chimeric and Variable Gene, Encodes an Extracellular Protein
RT Involved in Cellular Segregation in Streptococcus thermophilus.";
RL J. Bacteriol. 187:2737-2746(2005).
DR EMBL; CP000023; AAV59752.1; -; Genomic_DNA.
DR EMBL; AY730643; AAW82375.1; -; Genomic_DNA.
DR InterPro; IPR007921; CHAP.
DR InterPro; IPR009148; Siba.
DR Pfam; PF05257; CHAP; 1.
DR PRINTS; PR01852; SIBAPROTEIN.
DR PROSITE; PS0911; CHAP; 1.
KW Complete proteome.
SQ SEQUENCE 474 AA; 48142 MW; 1BAC6A9E0A0A200D CRC64;

Query Match 58.5%; Score 55; DB 2; Length 474;
Best Local Similarity 55.0%; Pred. No. 3.8;
Matches 11; Conservative 6; Mismatches 3; Indels 0; Gaps 0;

QY 1 QIASQDSKINNLTAQQQAAQ 20
DB 30 QIAATNNAISNLASQQEAAQ 49
||||: ::|||: |||: |||

RESULT 15
Q5M212 STRT1
ID Q5M212_STRT1 PRELIMINARY; PRT; 485 AA.
AC Q5M212;
DT 01-FEB-2005 (TREMBLrel. 29, Created)
DT 01-FEB-2005 (TREMBLrel. 29, Last sequence update)
DT 01-FEB-2005 (TREMBLrel. 29, Last annotation update)
DE Glucan binding protein.
GN Name=pcbB; OrderedLocustNames=str0022;
OS Streptococcus thermophilus (strain CNRZ 1066).
OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
OC Streptococcus.
OX NCBI_TaxID=299768;
RN [1]_
RP NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
RX PubMed=1543133; DOI=10.1038/nbt1034;
RA Bolotin A., Quinquis B., Renault P., Sorokin A., Ehrlich S.D.,
RA Kulakauskas S., Lapidus A., Goltsman E., Mazur M., Pusch G.D.,
RA Fonstein M., Overbeek R., Kyprides N., Purnelle B., Prozzi D.,
RA Ngui K., Masuy D., Hancy F., Burteau S., Boutry M., Delcour J.,
RA Goffeau A., Hols P.;
RT "Complete sequence and comparative genome analysis of the dairy
RT bacterium Streptococcus thermophilus.";
RL Nat. Biotechnol. 22:1554-1558(2004).
DR EMBL; CP000024; AAV61641.1; -; Genomic_DNA.
DR InterPro; IPR007921; CHAP.
DR InterPro; IPR009148; Siba.
DR Pfam; PF05257; CHAP; 1.
DR PRINTS; PR01852; SIBAPROTEIN.
DR PROSITE; PS0911; CHAP; 1.
KW Complete proteome.
SQ SEQUENCE 485 AA; 49165 MW; 8665E4B1FC6CF7CA CRC64;

Query Match 58.5%; Score 55; DB 2; Length 485;
Best Local Similarity 55.0%; Pred. No. 3.9;
Matches 11; Conservative 6; Mismatches 3; Indels 0; Gaps 0;

QY 1 QIASQDSKINNLTAQQQAAQ 20
DB 30 QIAATNNAISNLASQQEAAQ 49
||||: ::|||: |||: |||

Search completed: January 25, 2006, 19:13:39
Job time : 71.3 secs

```

Result No.	Score	Query Match	Length	DB	ID	Description	
1	74	78.7	210	2	US-09-222-938A-67	Sequence 67, Appl	
2	74	78.7	392	2	US-09-583-110-4374	Sequence 4374, Ap	
3	74	78.7	399	2	US-09-107-433-3230	Sequence 3230, Ap	
4	48	51.1	168	2	US-09-270-767-31753	Sequence 31753, A	
5	48	51.1	168	2	US-09-270-767-46970	Sequence 46970, A	
6	48	51.1	325	2	US-09-972-784-4	Sequence 4, Appli	
7	48	51.1	478	2	US-09-738-945-6	Sequence 6, Appli	
8	47	50.0	497	2	US-09-134-000C-5990	Sequence 5990, Ap	
9	45	47.9	282	2	US-09-605-703B-894	Sequence 894, App	
10	45	47.9	313	2	US-09-605-703B-892	Sequence 892, App	
11	44	46.8	328	2	US-09-270-767-32643	Sequence 32643, A	
12	44	46.8	328	2	US-09-270-767-47860	Sequence 47860, A	
13	44	46.8	523	2	US-09-489-039A-12885	Sequence 12885, A	
14	44	46.8	525	2	US-09-107-532A-5095	Sequence 5095, Ap	
15	43	45.7	243	2	US-09-328-352-7956	Sequence 7956, Ap	
16	42	44.7	207	2	US-09-489-039A-11667	Sequence 11667, A	
17	42	44.7	343	2	US-09-270-767-42868	Sequence 42868, A	
18	42	44.7	476	2	US-09-489-039A-10795	Sequence 10795, A	
19	42	44.7	506	2	US-09-134-001C-4733	Sequence 4733, Ap	
20	42	44.7	1220	2	US-09-540-236-3011	Sequence 3011, Ap	
21	42	44.7	1861	1	US-08-790-912-4	Sequence 4, Appli	
22	41	43.6	451	2	US-09-489-039A-11849	Sequence 11849, A	
23	41	43.6	666	2	US-09-248-796A-20656	Sequence 20656, A	
24	41	43.6	699	2	US-09-134-001C-4054	Sequence 4054, Ap	
25	41	43.6	731	1	US-08-696-944-20	Sequence 20, Appl	
26	41	43.6	1228	2	US-09-949-016-6805	Sequence 6805, Ap	
27	41	43.6	1236	2	US-09-949-016-10398	Sequence 10398, A	

```
; TYPE: PRT
; ORGANISM: Streptococcus pneumoniae
US-09-583-110-4374

Query Match      78.7%; Score 74; DB 2; Length 392;
Best Local Similarity 75.0%; Pred. No. 0.00031;
Matches 15; Conservative 4; Mismatches 1; Indels 0; Gaps 0;

QY 1 QIASQSKNNLTAAQQAAQ 20
   :||:||||:|||||
Db 33 KIAAQDNKISNLTAAQQEAQ 52

RESULT 3
US-09-107-433-3230
; Sequence 3230, Application US/09107433
; Patent No. 6800744
; GENERAL INFORMATION:
; APPLICANT: Lynn A Doucette-Stamm and David Bush
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID
; SEQUENCES RELATING TO STREPTOCOCCUS PNEUMONIAE FOR DIAGNOSTIC
; THERAPEUTICS
; NUMBER OF SEQUENCES: 5206
; CORRESPONDENCE ADDRESS:
; ADDRESSER: GENOME THERAPEUTICS CORPORATION
; STREET: 100 Beaver Street
; CITY: Waltham
; STATE: Massachusetts
; COUNTRY: USA
; ZIP: 02354
; COMPUTER READABLE FORM:
; MEDIUM TYPE: CD-ROM ISO9660
; COMPUTER: <Unknown>
; OPERATING SYSTEM: <Unknown>
; SOFTWARE: <Unknown>
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/107,433
; FILING DATE: 30-Jun-1998
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 60/ 085131
; FILING DATE: May 12, 1998
; APPLICATION NUMBER: 60/051553
; FILING DATE: July 2, 1997
; ATTORNEY/AGENT INFORMATION:
; NAME: Ariniello, Pamela Deneke
; REGISTRATION NUMBER: 40,489
; REFERENCE/DOCKET NUMBER: GTC-011
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (781)893-5007
; TELEFAX: (781)893-8277
; INFORMATION FOR SEQ ID NO: 3230:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 399 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; HYPOTHETICAL: YES
; ORIGINAL SOURCE:
; ORGANISM: Streptococcus pneumoniae
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (B) LOCATION 1...399
; SEQUENCE DESCRIPTION: SEQ ID NO: 3230:
US-09-107-433-3230

Query Match      78.7%; Score 74; DB 2; Length 399;
Best Local Similarity 75.0%; Pred. No. 0.00031;
Matches 15; Conservative 4; Mismatches 1; Indels 0; Gaps 0;

QY 1 QIASQSKNNLTAAQQAAQ 20
   :||:||||:|||||
Db 40 KIAAQDNKISNLTAAQQEAQ 59

; TYPE: PRT
; ORGANISM: Streptococcus pneumoniae
US-09-583-110-4374

Query Match      78.7%; Score 74; DB 2; Length 392;
Best Local Similarity 75.0%; Pred. No. 0.00031;
Matches 15; Conservative 4; Mismatches 1; Indels 0; Gaps 0;

QY 1 QIASQSKNNLTAAQQAAQ 20
   :||:||||:|||||
Db 33 KIAAQDNKISNLTAAQQEAQ 52

RESULT 4
US-09-270-767-31753
; Sequence 31753, Application US/09270767
; Patent No. 6703491
; GENERAL INFORMATION:
; APPLICANT: Homburger et al.
; TITLE OF INVENTION: Nucleic acids and proteins of Drosophila melanogaster
; FILE REFERENCE: File Reference: 7326-094
; CURRENT APPLICATION NUMBER: US/09/270,767
; CURRENT FILING DATE: 1999-03-17
; NUMBER OF SEQ ID NOS: 62517
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 31753
; LENGTH: 168
; TYPE: PRT
; ORGANISM: Drosophila melanogaster
US-09-270-767-31753

Query Match      51.1%; Score 48; DB 2; Length 168;
Best Local Similarity 62.5%; Pred. No. 2.3;
Matches 10; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

QY 5 QDSKNNLTAAQQAAQ 20
   :||:||||:|||||
Db 105 ENSKNNLAAQAQAQQ 120

RESULT 5
US-09-270-767-46970
; Sequence 46970, Application US/09270767
; Patent No. 6703491
; GENERAL INFORMATION:
; APPLICANT: Homburger et al.
; TITLE OF INVENTION: Nucleic acids and proteins of Drosophila melanogaster
; FILE REFERENCE: File Reference: 7326-094
; CURRENT APPLICATION NUMBER: US/09/270,767
; CURRENT FILING DATE: 1999-03-17
; NUMBER OF SEQ ID NOS: 62517
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 46970
; LENGTH: 168
; TYPE: PRT
; ORGANISM: Drosophila melanogaster
US-09-270-767-46970

Query Match      51.1%; Score 48; DB 2; Length 168;
Best Local Similarity 62.5%; Pred. No. 2.3;
Matches 10; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

QY 5 QDSKNNLTAAQQAAQ 20
   :||:||||:|||||
Db 105 ENSKNNLAAQAQAQQ 120

RESULT 6
US-09-972-784-4
; Sequence 4, Application US/09972784
; Patent No. 6566088
; GENERAL INFORMATION:
; APPLICANT: McKnight, Steven L.
; APPLICANT: Bruik, Richard K.
; TITLE OF INVENTION: Poly(1-4-Hydroxylases
; FILE REFERENCE: UTSD0871
; CURRENT APPLICATION NUMBER: US/09/972,784
; CURRENT FILING DATE: 2001-10-04
; NUMBER OF SEQ ID NOS: 5
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 4
; LENGTH: 325
; TYPE: PRT
; ORGANISM: Drosophila melanogaster
US-09-972-784-4
```

```

Query Match          51.1% ; Score 48 ; DB 2 ; Length 478 ;
Best Local Similarity 62.5% ; Pred. No. 7.7 ;
Matches 10 ; Conservative 2 ; Mismatches 4 ; Indels 0 ; Gaps 0 ;

QY      5 QDSKINNLTAAQQAAQ 20
        :| | | | | | | |
Db      415 ENSKTNLAAQAQAQ 430

RESULT 8
US-09-134-000C-5990
; Sequence 5990, Application US/09134000C
; Patent No. 6617156
; GENERAL INFORMATION:
; APPLICANT: Lynn Doucette-Stamm et al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO
; TITLE OF INVENTION: ENTEROCOCCUS FAECALIS FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 032796-032
; CURRENT APPLICATION NUMBER: US/09/134,000C
; CURRENT FILING DATE: 1998-08-13
; PRIOR APPLICATION NUMBER: US 60/055,778
; PRIOR FILING DATE: 1997-08-15
; NUMBER OF SEQ ID NOS: 6812
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 5990
; LENGTH: 497
; TYPE: PRT
; ORGANISM: Enterococcus faecalis
US-09-134-000C-5990

```

```

RESULT 10
US-09-605-703B-892
; Sequence 892, Application US/09605703B
; Patent No. 6962989
; GENERAL INFORMATION:
; APPLICANT: Pompejus, Markus
; APPLICANT: Kroger, Burkhard
; APPLICANT: Schroder, Hartwig
; APPLICANT: Zelder, Oskar
; APPLICANT: Haberhauer, Gregor
; TITLE OF INVENTION: CORYNEBACTERIUM GLUTAMICUM GENES ENCODING NOVEL
; FILE REFERENCE: BGI-129CP
; CURRENT APPLICATION NUMBER: US/09/605,703B
; CURRENT FILING DATE: 2000-06-27
; PRIOR APPLICATION NUMBER: 60/142,764
; PRIOR FILING DATE: 1999-07-08
; PRIOR APPLICATION NUMBER: 60/152,318
; PRIOR FILING DATE: 1999-09-03
; NUMBER OF SEQ ID NOS: 2934
; SEQ ID NO 892
; LENGTH: 313
; TYPE: PRT
; ORGANISM: Corynebacterium glutamicum
US-09-605-703B-892

Query Match          47.9%; Score 45; DB 2; Length 313;
Best Local Similarity 44.4%; Pred. No. 15;
Matches 8; Conservative 7; Mismatches 3; Indels 0; Gaps 0;

```

QY 3 ASQDSKINNLTAAQQAQ 20
 ::::: ||:||||
 Db 233 AAKQARVDLVAQQAQ 250

RESULT 11
 US-09-270-767-32643
 ; Sequence 32643, Application US/09270767
 ; Patent No. 6703491
 ; GENERAL INFORMATION:
 ; APPLICANT: Homburger et al.
 ; TITLE OF INVENTION: Nucleic acids and proteins of Drosophila melanogaster
 ; FILE REFERENCE: File Reference: 7326-094
 ; CURRENT APPLICATION NUMBER: US/09/270,767
 ; CURRENT FILING DATE: 1999-03-17
 ; NUMBER OF SEQ ID NOS: 62517
 ; SOFTWARE: PatentIn Ver. 2.0
 ; SEQ ID NO 32643
 ; LENGTH: 328
 ; TYPE: PRT
 ; ORGANISM: Drosophila melanogaster
 US-09-270-767-32643

Query Match 46.8%; Score 44; DB 2; Length 328;
 Best Local Similarity 45.0%; Pred. No. 23;
 Matches 9; Conservative 4; Mismatches 7; Indels 0; Gaps 0;

QY 1 QIASQDSKINNLTAAQQAQ 20
 ::||:|:|:|
 Db 201 KLAQDLRIERLDVQQLAQ 220

RESULT 12
 US-09-270-767-47860
 ; Sequence 47860, Application US/09270767
 ; Patent No. 6703491
 ; GENERAL INFORMATION:
 ; APPLICANT: Homburger et al.
 ; TITLE OF INVENTION: Nucleic acids and proteins of Drosophila melanogaster
 ; FILE REFERENCE: File Reference: 7326-094
 ; CURRENT APPLICATION NUMBER: US/09/270,767
 ; CURRENT FILING DATE: 1999-03-17
 ; NUMBER OF SEQ ID NOS: 62517
 ; SOFTWARE: PatentIn Ver. 2.0
 ; SEQ ID NO 47860
 ; LENGTH: 328
 ; TYPE: PRT
 ; ORGANISM: Drosophila melanogaster
 US-09-270-767-47860

Query Match 46.8%; Score 44; DB 2; Length 328;
 Best Local Similarity 45.0%; Pred. No. 23;
 Matches 9; Conservative 4; Mismatches 7; Indels 0; Gaps 0;

QY 1 QIASQDSKINNLTAAQQAQ 20
 ::||:|:|:|
 Db 201 KLAQDLRIERLDVQQLAQ 220

RESULT 13
 US-09-489-039A-12885
 ; Sequence 12885, Application US/09489039A
 ; Patent No. 6610836
 ; GENERAL INFORMATION:
 ; APPLICANT: Gary Breton et al.
 ; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO KLEBSIELLA
 ; FILE REFERENCE: PNEUMONIAE FOR DIAGNOSTICS AND THERAPEUTICS
 ; CURRENT APPLICATION NUMBER: US/09/489,039A
 ; CURRENT FILING DATE: 2000-01-27
 ; PRIOR APPLICATION NUMBER: US 60/117,747
 ; PRIOR FILING DATE: 1999-01-29

; NUMBER OF SEQ ID NOS: 14342
 ; SEQ ID NO 12885
 ; LENGTH: 523
 ; TYPE: PRT
 ; ORGANISM: Klebsiella pneumoniae
 US-09-489-039A-12885

Query Match 46.8%; Score 44; DB 2; Length 523;
 Best Local Similarity 50.0%; Pred. No. 39;
 Matches 9; Conservative 2; Mismatches 7; Indels 0; Gaps 0;

QY 3 ASQDSKINNLTAAQQAQ 20
 ::||:|:|:|
 Db 73 ASAEQKVQLTQQQQQTQ 90

RESULT 14
 US-09-107-532A-5095
 ; Sequence 5095, Application US/09107532A
 ; Patent No. 6583275
 ; GENERAL INFORMATION:
 ; APPLICANT: Lynn A Doucette-Stamm and David Bush
 ; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO
 ; ENTEROCOCCUS FAECIUM FOR DIAGNOSTICS AND THERAPEUTICS
 ; NUMBER OF SEQUENCES: 7310
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESS: GENOME THERAPEUTICS CORPORATION
 ; STREET: 100 Beaver Street
 ; CITY: Waltham
 ; STATE: Massachusetts
 ; COUNTRY: USA
 ; ZIP: 02354
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: CD-ROM ISO9660
 ; COMPUTER: PC
 ; OPERATING SYSTEM: <Unknown>
 ; SOFTWARE: ASCII
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/09/107,532A
 ; FILING DATE: 30-Jun-1998
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER: 60/085,598
 ; FILING DATE: 14 May 1998
 ; APPLICATION NUMBER: 60/051571
 ; FILING DATE: July 2, 1997
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: Ariniello, Pamela Deneke
 ; REGISTRATION NUMBER: 40,489
 ; REFERENCE/DOCKET NUMBER: GTC-012
 ; TELECOMMUNICATION INFORMATION:
 ; TELEPHONE: (781)893-5007
 ; TELEFAX: (781)893-8277
 ; INFORMATION FOR SEQ ID NO: 5095:
 ; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 525 amino acids
 ; TYPE: amino acid
 ; TOPOLOGY: linear
 ; MOLECULE TYPE: protein
 ; HYPOTHETICAL: YES
 ; ORIGINAL SOURCE:
 ; ORGANISM: Enterococcus faecium
 ; FEATURE:
 ; NAME/KEY: misc feature
 ; LOCATION: (B) LOCATION 1...525
 ; SEQUENCE DESCRIPTION: SEQ ID NO: 5095:

US-09-107-532A-5095
 ;

Query Match 46.8%; Score 44; DB 2; Length 525;
 Best Local Similarity 55.0%; Pred. No. 40;
 Matches 11; Conservative 1; Mismatches 8; Indels 0; Gaps 0;

QY 1 QIASQDSKINNLTAAQQAQ 20
 ::||:|:|:|

```

RESULT 15
US-09-328-352-7956
; Sequence 7956, Application US/09328352
; Patent No. 6562958
; GENERAL INFORMATION:
; APPLICANT: Gary L. Breton et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO ACINETOBACTER
; TITLE OF INVENTION: BAUMANNII FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: GTC99-03PA
; CURRENT APPLICATION NUMBER: US/09/328,352
; CURRENT FILING DATE: 1999-06-04
; NUMBER OF SEQ ID NOS: 8252
; SEQ ID NO 7956
; LENGTH: 243
; TYPE: PRT
; ORGANISM: Acinetobacter baumannii
US-09-328-352-7956

```

```
Query Match      45.7%; Score 43; DB 2; Length 243;
Best Local Similarity 47.4%; Pred. No. 23;
Matches 9; Conservative 4; Mismatches 6; Indels 0; Gaps 0;
```

Qy 1 QIASQDSKINNLTAAQQAA 19
 | : | : | : | : |
Db 61 QLAEQSKVQNQSATTOAA 79

Search completed: January 25, 2006, 19:18:14
Job time : 17.6 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2006 Compugen Ltd.

OM protein - protein search, using sw model

Run on: January 25, 2006, 19:13:51 ; Search time 60 Seconds
(without alignments)
139.276 Million cell updates/sec

Title: US-10-797-821-6
Perfect score: 94
Sequence: 1 QIASQDSKINNLTAAQQAAQ 20

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1867569 seqs, 417829326 residues

Total number of hits satisfying chosen parameters: 1867569

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published Applications AA Main:*
1: /cgn2_6/prodata1/pubpaa/US07_PUBCOMB.pep:*
2: /cgn2_6/prodata1/pubpaa/US08_PUBCOMB.pep:*
3: /cgn2_6/prodata1/pubpaa/US09_PUBCOMB.pep:*
4: /cgn2_6/prodata1/pubpaa/US10A_PUBCOMB.pep:*
5: /cgn2_6/prodata1/pubpaa/US10B_PUBCOMB.pep:*
6: /cgn2_6/prodata1/pubpaa/US11_PUBCOMB.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	94	100.0	20	4	US-10-383-930-6
2	94	100.0	20	5	US-10-797-821-6
3	94	100.0	431	4	US-10-383-930-29
4	94	100.0	431	4	US-10-383-930-30
5	94	100.0	431	4	US-10-383-930-33
6	94	100.0	431	5	US-10-797-821-29
7	94	100.0	431	5	US-10-797-821-30
8	94	100.0	431	5	US-10-797-821-33
9	94	100.0	432	4	US-10-383-930-31
10	94	100.0	432	4	US-10-383-930-32
11	94	100.0	432	5	US-10-797-821-31
12	94	100.0	432	5	US-10-797-821-32
13	77	81.9	20	4	US-10-383-930-7
14	77	81.9	20	5	US-10-797-821-7
15	74	78.7	210	4	US-10-154-251-67
16	74	78.7	392	5	US-10-472-928-4652
17	74	78.7	399	5	US-10-617-320-3230
18	62	66.0	398	5	US-10-474-792-600
19	48	51.1	325	6	US-11-097-143-117
20	47	50.0	618	6	US-11-097-143-23781
21	46	48.9	115	4	US-10-425-115-288800
22	46	48.9	199	4	US-10-425-114-40322
23	45	47.9	371	3	US-09-738-626-5471
24	45	47.9	1059	4	US-10-437-963-147607
25	45	47.9	1112	4	US-10-437-963-147651
26	44	46.8	341	6	US-11-097-143-25374
27	44	46.8	524	4	US-10-282-122A-57658

28 44 46.8 1299 4 US-10-437-963-147640 Sequence 147640,
29 43 45.7 212 4 US-10-767-701-53332 Sequence 53332, A
30 43 45.7 243 5 US-10-805-684-18 Sequence 18, Appl
31 43 45.7 302 4 US-10-108-260A-2659 Sequence 2659, Ap
32 43 45.7 562 4 US-10-425-114-52777 Sequence 52777, A
33 43 45.7 604 4 US-10-425-115-286982 Sequence 286982,
34 43 45.7 668 5 US-10-805-684-40 Sequence 40, Appl
35 42 44.7 235 5 US-10-479-925-6 Sequence 6, Appl
36 42 44.7 235 5 US-10-732-923-14656 Sequence 14656, A
37 42 44.7 295 4 US-10-424-599-149551 Sequence 149551,
38 42 44.7 308 6 US-11-097-143-17952 Sequence 17952, A
39 42 44.7 499 4 US-10-424-599-249978 Sequence 249978,
40 42 44.7 506 4 US-10-724-972A-6245 Sequence 6245, Ap
41 42 44.7 653 4 US-10-437-963-186821 Sequence 186821,
42 42 44.7 722 4 US-10-282-122A-51562 Sequence 51562, A
43 42 44.7 1164 5 US-10-732-923-3344 Sequence 3344, Ap
44 42 44.7 1199 4 US-10-282-122A-63269 Sequence 63269, A
45 42 44.7 1857 6 US-11-097-143-20154 Sequence 20154, A

ALIGNMENTS

RESULT 1
US-10-383-930-6
; Sequence 6, Application US/10383930
; Publication NO. US20040127400A1
; GENERAL INFORMATION:
; APPLICANT: Smith, Daniel J
; APPLICANT: Taubman, Martin A
; TITLE OF INVENTION: Immunogenicity of Glucan Binding Protein
; FILE REFERENCE: 25669-018
; CURRENT APPLICATION NUMBER: US/10/383,930
; CURRENT FILING DATE: 2003-03-07
; PRIOR APPLICATION NUMBER: 60/402,483
; PRIOR FILING DATE: 2002-08-08
; PRIOR APPLICATION NUMBER: 60/363,209
; PRIOR FILING DATE: 2002-03-07
; NUMBER OF SEQ ID NOS: 41
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 6
; LENGTH: 20
; TYPE: PRT
; ORGANISM: Streptococcus mutans
US-10-383-930-6

Query Match 100.0%; Score 94; DB 4; Length 20;
Best Local Similarity 100.0%; Pred.No. 7.1e-08;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 QIASQDSKINNLTAAQQAAQ 20
Db 1 QIASQDSKINNLTAAQQAAQ 20

RESULT 2
US-10-797-821-6
; Sequence 6, Application US/10797821
; Publication No. US20050031633A1
; GENERAL INFORMATION:
; APPLICANT: Smith, Daniel J.
; APPLICANT: Taubman, Martin A.
; TITLE OF INVENTION: Glucan Binding Protein and Glycosyltransferase Immunogens
; FILE REFERENCE: 25669-020
; CURRENT APPLICATION NUMBER: US/10/797,821
; CURRENT FILING DATE: 2004-03-09
; PRIOR APPLICATION NUMBER: 10/383,930
; PRIOR FILING DATE: 2003-03-07
; PRIOR APPLICATION NUMBER: 60/363,209
; PRIOR FILING DATE: 2002-03-07
; PRIOR APPLICATION NUMBER: 60/402,483
; PRIOR FILING DATE: 2002-08-08
; PRIOR APPLICATION NUMBER: 09/290,049

; PRIOR FILING DATE: 1999-04-12
; PRIOR APPLICATION NUMBER: 60/081,550
; PRIOR FILING DATE: 1998-04-13
; PRIOR APPLICATION NUMBER: 60/115,142
; PRIOR FILING DATE: 1999-01-08
; NUMBER OF SEQ ID NOS: 45
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 6
; LENGTH: 20
; TYPE: PRT
; ORGANISM: Artificial
; FEATURE:
; OTHER INFORMATION: GbpB peptide
US-10-797-821-6

Query Match 100.0%; Score 94; DB 5; Length 20;
Best Local Similarity 100.0%; Pred. No. 7.1e-08;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 QIASQDSKINNLTAAQQAAQ 20
|||
Db 1 QIASQDSKINNLTAAQQAAQ 20

RESULT 3

US-10-383-930-29
; Sequence 29, Application US/10383930
; Publication No. US20040127400A1
; GENERAL INFORMATION:
; APPLICANT: Smith, Daniel J
; APPLICANT: Taubman, Martin A
; TITLE OF INVENTION: Immunogenicity of Glucan Binding Protein
; FILE REFERENCE: 25669-018
; CURRENT APPLICATION NUMBER: US/10/383,930
; CURRENT FILING DATE: 2003-03-07
; PRIOR APPLICATION NUMBER: 60/402,483
; PRIOR FILING DATE: 2002-08-08
; PRIOR APPLICATION NUMBER: 60/363,209
; PRIOR FILING DATE: 2002-03-07
; NUMBER OF SEQ ID NOS: 41
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 29
; LENGTH: 431
; TYPE: PRT
; ORGANISM: Streptococcus mutans
US-10-383-930-29

Query Match 100.0%; Score 94; DB 4; Length 431;
Best Local Similarity 100.0%; Pred. No. 2.1e-06;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 QIASQDSKINNLTAAQQAAQ 20
|||
Db 33 QIASQDSKINNLTAAQQAAQ 52

RESULT 4

US-10-383-930-30
; Sequence 30, Application US/10383930
; Publication No. US20040127400A1
; GENERAL INFORMATION:
; APPLICANT: Smith, Daniel J
; APPLICANT: Taubman, Martin A
; TITLE OF INVENTION: Immunogenicity of Glucan Binding Protein
; FILE REFERENCE: 25669-018
; CURRENT APPLICATION NUMBER: US/10/383,930
; CURRENT FILING DATE: 2003-03-07
; PRIOR APPLICATION NUMBER: 60/402,483
; PRIOR FILING DATE: 2002-08-08
; PRIOR APPLICATION NUMBER: 60/363,209
; PRIOR FILING DATE: 2002-03-07
; NUMBER OF SEQ ID NOS: 41
; SOFTWARE: PatentIn version 3.2

; SEQ ID NO 30
; LENGTH: 431
; TYPE: PRT
; ORGANISM: Streptococcus mutans
US-10-383-930-30

Query Match 100.0%; Score 94; DB 4; Length 431;
Best Local Similarity 100.0%; Pred. No. 2.1e-06;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 QIASQDSKINNLTAAQQAAQ 20
|||
Db 33 QIASQDSKINNLTAAQQAAQ 52

RESULT 5

US-10-383-930-33
; Sequence 33, Application US/10383930
; Publication No. US20040127400A1
; GENERAL INFORMATION:
; APPLICANT: Smith, Daniel J
; APPLICANT: Taubman, Martin A
; TITLE OF INVENTION: Immunogenicity of Glucan Binding Protein
; FILE REFERENCE: 25669-018
; CURRENT APPLICATION NUMBER: US/10/383,930
; CURRENT FILING DATE: 2003-03-07
; PRIOR APPLICATION NUMBER: 60/402,483
; PRIOR FILING DATE: 2002-08-08
; PRIOR APPLICATION NUMBER: 60/363,209
; PRIOR FILING DATE: 2002-03-07
; NUMBER OF SEQ ID NOS: 41
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 33
; LENGTH: 431
; TYPE: PRT
; ORGANISM: Streptococcus mutans
US-10-383-930-33

Query Match 100.0%; Score 94; DB 4; Length 431;
Best Local Similarity 100.0%; Pred. No. 2.1e-06;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 QIASQDSKINNLTAAQQAAQ 20
|||
Db 33 QIASQDSKINNLTAAQQAAQ 52

RESULT 6

US-10-797-821-29
; Sequence 29, Application US/10797821
; Publication No. US20050031633A1
; GENERAL INFORMATION:
; APPLICANT: Smith, Daniel J.
; APPLICANT: Taubman, Martin A.
; TITLE OF INVENTION: Glucan Binding Protein and Glycosyltransferase Immunogens
; FILE REFERENCE: 25669-020
; CURRENT APPLICATION NUMBER: US/10/797,821
; CURRENT FILING DATE: 2004-03-09
; PRIOR APPLICATION NUMBER: 10/383,930
; PRIOR FILING DATE: 2003-03-07
; PRIOR APPLICATION NUMBER: 60/363,209
; PRIOR FILING DATE: 2002-03-07
; PRIOR APPLICATION NUMBER: 60/402,483
; PRIOR FILING DATE: 2002-08-08
; PRIOR APPLICATION NUMBER: 09/290,049
; PRIOR FILING DATE: 1999-04-12
; PRIOR APPLICATION NUMBER: 60/081,550
; PRIOR FILING DATE: 1998-04-13
; PRIOR APPLICATION NUMBER: 60/115,142
; PRIOR FILING DATE: 1999-01-08
; NUMBER OF SEQ ID NOS: 45
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 29

```
; LENGTH: 431
; TYPE: PRT
; ORGANISM: Streptococcus mutans
US-10-797-821-29

Query Match      100.0%; Score 94; DB 5; Length 431;
Best Local Similarity 100.0%; Pred. No. 2.1e-06;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 QIASQDSKINNLTAAQQAAQ 20
   |||||
Db 33 QIASQDSKINNLTAAQQAAQ 52

RESULT 7
US-10-797-821-30
; Sequence 30, Application US/10797821
; Publication No. US20050031633A1
; GENERAL INFORMATION:
; APPLICANT: Smith, Daniel J.
; APPLICANT: Taubman, Martin A.
; TITLE OF INVENTION: Glucan Binding Protein and Glycosyltransferase Immunogens
; FILE REFERENCE: 25669-020
; CURRENT APPLICATION NUMBER: US/10/797,821
; CURRENT FILING DATE: 2004-03-09
; PRIOR APPLICATION NUMBER: 10/383,930
; PRIOR FILING DATE: 2003-03-07
; PRIOR APPLICATION NUMBER: 60/363,209
; PRIOR FILING DATE: 2002-03-07
; PRIOR APPLICATION NUMBER: 60/402,483
; PRIOR FILING DATE: 2002-08-08
; PRIOR APPLICATION NUMBER: 09/290,049
; PRIOR FILING DATE: 1999-04-12
; PRIOR APPLICATION NUMBER: 60/081,550
; PRIOR FILING DATE: 1998-04-13
; PRIOR APPLICATION NUMBER: 60/115,142
; PRIOR FILING DATE: 1999-01-08
; NUMBER OF SEQ ID NOS: 45
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 30
; LENGTH: 431
; TYPE: PRT
; ORGANISM: Streptococcus mutans
US-10-797-821-30

Query Match      100.0%; Score 94; DB 5; Length 431;
Best Local Similarity 100.0%; Pred. No. 2.1e-06;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 QIASQDSKINNLTAAQQAAQ 20
   |||||
Db 33 QIASQDSKINNLTAAQQAAQ 52

RESULT 8
US-10-797-821-33
; Sequence 33, Application US/10797821
; Publication No. US20050031633A1
; GENERAL INFORMATION:
; APPLICANT: Smith, Daniel J.
; APPLICANT: Taubman, Martin A.
; TITLE OF INVENTION: Glucan Binding Protein and Glycosyltransferase Immunogens
; FILE REFERENCE: 25669-020
; CURRENT APPLICATION NUMBER: US/10/797,821
; CURRENT FILING DATE: 2004-03-09
; PRIOR APPLICATION NUMBER: 10/383,930
; PRIOR FILING DATE: 2003-03-07
; PRIOR APPLICATION NUMBER: 60/363,209
; PRIOR FILING DATE: 2002-03-07
; PRIOR APPLICATION NUMBER: 60/402,483
; PRIOR FILING DATE: 2002-08-08
; PRIOR APPLICATION NUMBER: 09/290,049
; PRIOR FILING DATE: 1999-04-12
```

```
; PRIOR APPLICATION NUMBER: 60/081,550
; PRIOR FILING DATE: 1998-04-13
; PRIOR APPLICATION NUMBER: 60/115,142
; PRIOR FILING DATE: 1999-01-08
; NUMBER OF SEQ ID NOS: 45
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 33
; LENGTH: 431
; TYPE: PRT
; ORGANISM: Streptococcus mutans
US-10-797-821-33

Query Match      100.0%; Score 94; DB 5; Length 431;
Best Local Similarity 100.0%; Pred. No. 2.1e-06;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 QIASQDSKINNLTAAQQAAQ 20
   |||||
Db 33 QIASQDSKINNLTAAQQAAQ 52

RESULT 9
US-10-383-930-31
; Sequence 31, Application US/10383930
; Publication No. US20040127400A1
; GENERAL INFORMATION:
; APPLICANT: Smith, Daniel J.
; APPLICANT: Taubman, Martin A.
; TITLE OF INVENTION: Immunogenicity of Glucan Binding Protein
; FILE REFERENCE: 25669-018
; CURRENT APPLICATION NUMBER: US/10/383,930
; CURRENT FILING DATE: 2003-03-07
; PRIOR APPLICATION NUMBER: 60/402,483
; PRIOR FILING DATE: 2002-08-08
; PRIOR APPLICATION NUMBER: 60/363,209
; PRIOR FILING DATE: 2002-03-07
; NUMBER OF SEQ ID NOS: 41
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 31
; LENGTH: 432
; TYPE: PRT
; ORGANISM: Streptococcus mutans
US-10-383-930-31

Query Match      100.0%; Score 94; DB 4; Length 432;
Best Local Similarity 100.0%; Pred. No. 2.1e-06;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 QIASQDSKINNLTAAQQAAQ 20
   |||||
Db 33 QIASQDSKINNLTAAQQAAQ 52

RESULT 10
US-10-383-930-32
; Sequence 32, Application US/10383930
; Publication No. US20040127400A1
; GENERAL INFORMATION:
; APPLICANT: Smith, Daniel J.
; APPLICANT: Taubman, Martin A.
; TITLE OF INVENTION: Immunogenicity of Glucan Binding Protein
; FILE REFERENCE: 25669-018
; CURRENT APPLICATION NUMBER: US/10/383,930
; CURRENT FILING DATE: 2003-03-07
; PRIOR APPLICATION NUMBER: 60/402,483
; PRIOR FILING DATE: 2002-08-08
; PRIOR APPLICATION NUMBER: 60/363,209
; PRIOR FILING DATE: 2002-03-07
; NUMBER OF SEQ ID NOS: 41
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 32
; LENGTH: 432
; TYPE: PRT
```

```
; ORGANISM: Streptococcus mutans
US-10-383-930-32

Query Match      100.0%; Score 94; DB 4; Length 432;
Best Local Similarity 100.0%; Pred. No. 2.1e-06;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 QIASQDSKINNLTAQQQAAQ 20
   |||||
Db 33 QIASQDSKINNLTAQQQAAQ 52

RESULT 11
US-10-797-821-31
; Sequence 31, Application US/10797821
; Publication No. US20050031633A1
; GENERAL INFORMATION:
; APPLICANT: Smith, Daniel J.
; APPLICANT: Taubman, Martin A.
; TITLE OF INVENTION: Glucan Binding Protein and Glycosyltransferase Immunogens
; FILE REFERENCE: 25669-020
; CURRENT APPLICATION NUMBER: US/10/797,821
; CURRENT FILING DATE: 2004-03-09
; PRIOR APPLICATION NUMBER: 10/383,930
; PRIOR FILING DATE: 2003-03-07
; PRIOR APPLICATION NUMBER: 60/363,209
; PRIOR FILING DATE: 2002-03-07
; PRIOR APPLICATION NUMBER: 60/402,483
; PRIOR FILING DATE: 1999-04-12
; PRIOR APPLICATION NUMBER: 60/081,550
; PRIOR FILING DATE: 1998-04-13
; PRIOR APPLICATION NUMBER: 60/115,142
; PRIOR FILING DATE: 1999-01-08
; NUMBER OF SEQ ID NOS: 45
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 31
; LENGTH: 432
; TYPE: PRT
; ORGANISM: Streptococcus mutans
US-10-797-821-31

Query Match      100.0%; Score 94; DB 5; Length 432;
Best Local Similarity 100.0%; Pred. No. 2.1e-06;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 QIASQDSKINNLTAQQQAAQ 20
   |||||
Db 33 QIASQDSKINNLTAQQQAAQ 52

RESULT 12
US-10-797-821-32
; Sequence 32, Application US/10797821
; Publication No. US20050031633A1
; GENERAL INFORMATION:
; APPLICANT: Smith, Daniel J.
; APPLICANT: Taubman, Martin A.
; TITLE OF INVENTION: Glucan Binding Protein and Glycosyltransferase Immunogens
; FILE REFERENCE: 25669-020
; CURRENT APPLICATION NUMBER: US/10/797,821
; CURRENT FILING DATE: 2004-03-09
; PRIOR APPLICATION NUMBER: 10/383,930
; PRIOR FILING DATE: 2003-03-07
; PRIOR APPLICATION NUMBER: 60/363,209
; PRIOR FILING DATE: 2002-03-07
; PRIOR APPLICATION NUMBER: 60/402,483
; PRIOR FILING DATE: 2002-08-08
; PRIOR APPLICATION NUMBER: 09/290,049
; PRIOR FILING DATE: 1999-04-12
; PRIOR APPLICATION NUMBER: 60/081,550
; PRIOR FILING DATE: 1998-04-13

; ORGANISM: Streptococcus mutans
US-10-797-821-32

Query Match      100.0%; Score 94; DB 5; Length 432;
Best Local Similarity 100.0%; Pred. No. 2.1e-06;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 QIASQDSKINNLTAQQQAAQ 20
   |||||
Db 33 QIASQDSKINNLTAQQQAAQ 52

RESULT 13
US-10-383-930-7
; Sequence 7, Application US/10383930
; Publication No. US20040127400A1
; GENERAL INFORMATION:
; APPLICANT: Smith, Daniel J.
; APPLICANT: Taubman, Martin A.
; TITLE OF INVENTION: Immunogenicity of Glucan Binding Protein
; FILE REFERENCE: 25669-018
; CURRENT APPLICATION NUMBER: US/10/383,930
; CURRENT FILING DATE: 2003-03-07
; PRIOR APPLICATION NUMBER: 60/402,483
; PRIOR FILING DATE: 2002-08-08
; PRIOR APPLICATION NUMBER: 60/363,209
; PRIOR FILING DATE: 2002-03-07
; NUMBER OF SEQ ID NOS: 41
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 7
; LENGTH: 20
; TYPE: PRT
; ORGANISM: Streptococcus mutans
US-10-383-930-7

Query Match      81.9%; Score 77; DB 4; Length 20;
Best Local Similarity 100.0%; Pred. No. 3.4e-05;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 5 QDSKINNLTAQQQAAQ 20
   |||||
Db 1 QDSKINNLTAQQQAAQ 16

RESULT 14
US-10-797-821-7
; Sequence 7, Application US/10797821
; Publication No. US20050031633A1
; GENERAL INFORMATION:
; APPLICANT: Smith, Daniel J.
; APPLICANT: Taubman, Martin A.
; TITLE OF INVENTION: Glucan Binding Protein and Glycosyltransferase Immunogens
; FILE REFERENCE: 25669-020
; CURRENT APPLICATION NUMBER: US/10/797,821
; CURRENT FILING DATE: 2004-03-09
; PRIOR APPLICATION NUMBER: 10/383,930
; PRIOR FILING DATE: 2003-03-07
; PRIOR APPLICATION NUMBER: 60/363,209
; PRIOR FILING DATE: 2002-03-07
; PRIOR APPLICATION NUMBER: 60/402,483
; PRIOR FILING DATE: 2002-08-08
; PRIOR APPLICATION NUMBER: 09/290,049
; PRIOR FILING DATE: 1999-04-12
; PRIOR APPLICATION NUMBER: 60/081,550
; PRIOR FILING DATE: 1998-04-13
; PRIOR APPLICATION NUMBER: 60/115,142
```


THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2006 Compugen Ltd.

OM protein - protein search, using sw model

Run on: January 25, 2006, 19:15:26 ; Search time 5.5 Seconds
(without alignments)
39.378 Million cell updates/sec

Title: US-10-797-821-6
Perfect score: 94
Sequence: 1 QIASQDSKINNLTAAQQAAQ 20

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 75621 seqs, 10829074 residues

Total number of hits satisfying chosen parameters: 75621

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

- Database : Published Applications_AA_New.*
- 1: /cgn2_6/prodata/2/pubpaa/US08_NEW_PUB.pep.*
 - 2: /cgn2_6/prodata/2/pubpaa/US06_NEW_PUB.pep.*
 - 3: /cgn2_6/prodata/2/pubpaa/US07_NEW_PUB.pep.*
 - 4: /cgn2_6/prodata/2/pubpaa/PCT_NEW_PUB.pep.*
 - 5: /cgn2_6/prodata/2/pubpaa/US05_NEW_PUB.pep.*
 - 6: /cgn2_6/prodata/2/pubpaa/US10_NEW_PUB.pep.*
 - 7: /cgn2_6/prodata/2/pubpaa/US11_NEW_PUB.pep.*
 - 8: /cgn2_6/prodata/2/pubpaa/US60_NEW_PUB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	94	100.0	431	7	US-11-052-554A-210
2	62	66.0	398	7	US-11-052-554A-252
3	40	42.6	205	6	US-10-467-657-4336
4	40	42.6	1404	6	US-10-878-556A-169
5	39	41.5	489	6	US-10-537-075-11
6	38	40.4	459	6	US-10-467-657-8726
7	38	40.4	731	6	US-10-613-744-10
8	38	40.4	1145	6	US-10-793-626-1432
9	37	39.4	261	6	US-10-821-234-1382
10	37	39.4	860	7	US-11-019-711-59
11	37	39.4	1050	6	US-10-523-477-12
12	37	39.4	1050	6	US-10-770-726-47
13	36	38.3	445	6	US-10-957-887B-174
14	36	38.3	89	7	US-11-145-631-16
15	36	38.3	228	7	US-11-074-176-136
16	36	38.3	352	6	US-10-793-626-216
17	36	38.3	646	6	US-10-793-626-676
18	36	38.3	665	6	US-10-873-528-127
19	36	38.3	687	6	US-10-467-657-1300
20	36	38.3	766	7	US-11-147-047-27
21	36	38.3	850	7	US-11-037-243-108
22	36	38.3	1242	6	US-10-517-939-308
23	36	38.3	1448	6	US-10-485-517-212
24	36	38.3	1493	7	US-11-004-057-4
25	36	38.3	1960	7	US-11-069-834-48

26	36	38.3	1960	7	US-11-069-834-50	Sequence 50, Appl
27	35	37.2	127	6	US-10-793-626-2358	Sequence 2358, Ap
28	35	37.2	315	7	US-11-024-959-482	Sequence 482, App
29	35	37.2	364	6	US-10-877-346-64	Sequence 64, Appl
30	35	37.2	455	6	US-10-508-263-66	Sequence 66, Appl
31	35	37.2	507	7	US-11-052-554A-216	Sequence 216, App
32	35	37.2	537	6	US-10-873-528-66	Sequence 66, Appl
33	35	37.2	612	6	US-10-467-657-3988	Sequence 3988, Ap
34	35	37.2	628	6	US-10-467-657-2484	Sequence 2484, Ap
35	35	37.2	628	6	US-10-995-561-1002	Sequence 1002, Ap
36	35	37.2	716	7	US-11-194-890-20	Sequence 20, Appl
37	35	37.2	863	7	US-11-097-749-2	Sequence 2, Appl
38	35	37.2	915	6	US-10-821-234-1514	Sequence 1514, Ap
39	35	37.2	915	6	US-10-995-561-1003	Sequence 1003, Ap
40	35	37.2	917	6	US-10-995-561-1000	Sequence 1000, Ap
41	35	37.2	940	6	US-10-995-561-1004	Sequence 1004, Ap
42	35	37.2	969	6	US-10-995-561-1001	Sequence 1001, Ap
43	35	37.2	971	6	US-10-995-561-998	Sequence 998, App
44	35	37.2	994	6	US-10-995-561-997	Sequence 997, App
45	35	37.2	1167	6	US-10-601-368-18	Sequence 18, Appl

ALIGNMENTS

RESULT 1
US-11-052-554A-210
; Sequence 210, Application US/11052554A
; Publication No. US20050288866A1
; GENERAL INFORMATION:
; APPLICANT: Sachdeva, et al.
; TITLE OF INVENTION: COMPUTATIONAL METHOD FOR IDENTIFYING ADHESIN AND ADHESIN-LIKE
; FILE REFERENCE: 30853/40359A
; CURRENT APPLICATION NUMBER: US/11/052,554A
; CURRENT FILING DATE: 2005-02-07
; PRIOR APPLICATION NUMBER: US 60/589,227
; PRIOR FILING DATE: 2004-07-20
; PRIOR APPLICATION NUMBER: IN 173/DEL/2004
; PRIOR FILING DATE: 2004-02-06
; NUMBER OF SEQ ID NOS: 763
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 210
; LENGTH: 431
; TYPE: PRT
; ORGANISM: Streptococcus mutans UA159
US-11-052-554A-210

Query Match 100.0%; Score 94; DB 7; Length 431;
Best Local Similarity 100.0%; Pred. No. 3.2e-08;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 QIASQDSKINNLTAAQQAAQ 20
DB 33 QIASQDSKINNLTAAQQAAQ 52

RESULT 2
US-11-052-554A-252
; Sequence 252, Application US/11052554A
; Publication No. US20050288866A1
; GENERAL INFORMATION:
; APPLICANT: Sachdeva, et al.
; TITLE OF INVENTION: COMPUTATIONAL METHOD FOR IDENTIFYING ADHESIN AND ADHESIN-LIKE
; FILE REFERENCE: 30853/40359A
; CURRENT APPLICATION NUMBER: US/11/052,554A
; CURRENT FILING DATE: 2005-02-07
; PRIOR APPLICATION NUMBER: US 60/589,227
; PRIOR FILING DATE: 2004-07-20
; PRIOR APPLICATION NUMBER: IN 173/DEL/2004
; PRIOR FILING DATE: 2004-02-06
; NUMBER OF SEQ ID NOS: 763

; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 252
; LENGTH: 398
; TYPE: PRT
; ORGANISM: Streptococcus pyogenes MGAS8232
US-11-052-554A-252

Query Match 66.0%; Score 62; DB 7; Length 398;
Best Local Similarity 65.0%; Pred. No. 0.0049;
Matches 13; Conservative 4; Mismatches 3; Indels 0; Gaps 0;

QY 1 QIASQDSKINNLTAAQQAQ 20
Db 30 KIAQDSIISNLTTEQKAQ 49

RESULT 3

US-10-467-657-4336
; Sequence 4336, Application US/10467657
; Publication No. US20050260581A1
; GENERAL INFORMATION:
; APPLICANT: CHIRON SpA
; APPLICANT: FONTANA Maria Rita
; APPLICANT: PIZZA Mariagrazia
; APPLICANT: MASIGNANI Vega
; APPLICANT: MONACI Elisabetta
; TITLE OF INVENTION: GONOCOCCAL PROTEINS AND NUCLEIC ACIDS
; FILE REFERENCE:
; CURRENT APPLICATION NUMBER: US/10/467,657
; CURRENT FILING DATE: 2003-08-11
; PRIOR APPLICATION NUMBER: GB-0103424.8
; PRIOR FILING DATE: 2001-02-12
; NUMBER OF SEQ ID NOS: 9218
; SOFTWARE: SeqWin99, version 1.04
; SEQ ID NO 4336
; LENGTH: 205
; TYPE: PRT
; ORGANISM: Neisseria gonorrhoeae
US-10-467-657-4336

Query Match 42.6%; Score 40; DB 6; Length 205;
Best Local Similarity 52.9%; Pred. No. 9;
Matches 9; Conservative 3; Mismatches 5; Indels 0; Gaps 0;

QY 4 SQDSKINNLTAAQQAQ 20
Db 163 SRDLNTRSLRAKQAQ 179

RESULT 4

US-10-878-556A-169
; Sequence 169, Application US/10878556A
; Publication No. US20050266399A1
; GENERAL INFORMATION:
; APPLICANT: Hoffmann La-Roche Inc.
; TITLE OF INVENTION: HCV regulated protein expression
; FILE REFERENCE: 21762
; CURRENT APPLICATION NUMBER: US/10/878,556A
; CURRENT FILING DATE: 2004-06-28
; NUMBER OF SEQ ID NOS: 199
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 169
; LENGTH: 1404
; TYPE: PRT
; ORGANISM: Homo sapiens
; PUBLICATION INFORMATION:
; DATABASE ACCESSION NUMBER: humangp/chr12-q14221
; DATABASE ENTRY DATE: 2003-04-22
US-10-878-556A-169

Query Match 42.6%; Score 40; DB 6; Length 1404;
Best Local Similarity 45.0%; Pred. No. 82;
Matches 9; Conservative 3; Mismatches 8; Indels 0; Gaps 0;

QY 1 QIASQDSKINNLTAAQQAQ 20
Db 510 QLGDKDQKIQLNEALLQSK 529

RESULT 5

US-10-537-075-11
; Sequence 11, Application US/10537075
; Publication No. US20060014291A1
; GENERAL INFORMATION:
; APPLICANT: Kesseler, Maria
; APPLICANT: Zelinski, Thomas
; APPLICANT: Hauer, Bernhard
; TITLE OF INVENTION: L-RHAMNOSE-INDUCIBLE EXPRESSION SYSTEMS
; FILE REFERENCE: 12810-00091-US
; CURRENT APPLICATION NUMBER: US/10/537,075
; CURRENT FILING DATE: 2005-06-01
; PRIOR APPLICATION NUMBER: PCT/EP2003/013367
; PRIOR FILING DATE: 2003-11-27
; PRIOR APPLICATION NUMBER: DE 102 56 381.0
; PRIOR FILING DATE: 2002-12-02
; NUMBER OF SEQ ID NOS: 19
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 11
; LENGTH: 489
; TYPE: PRT
; ORGANISM: Escherichia coli
US-10-537-075-11

Query Match 41.5%; Score 39; DB 6; Length 489;
Best Local Similarity 50.0%; Pred. No. 35;
Matches 8; Conservative 3; Mismatches 5; Indels 0; Gaps 0;

QY 2 IASQDSKINNLTAAQQAQ 17
Db 100 VAYRDSRTNGLMAQAQ 115

RESULT 6

US-10-467-657-8726
; Sequence 8726, Application US/10467657
; Publication No. US20050260581A1
; GENERAL INFORMATION:
; APPLICANT: CHIRON SpA
; APPLICANT: FONTANA Maria Rita
; APPLICANT: PIZZA Mariagrazia
; APPLICANT: MASIGNANI Vega
; APPLICANT: MONACI Elisabetta
; TITLE OF INVENTION: GONOCOCCAL PROTEINS AND NUCLEIC ACIDS
; FILE REFERENCE:
; CURRENT APPLICATION NUMBER: US/10/467,657
; CURRENT FILING DATE: 2003-08-11
; PRIOR APPLICATION NUMBER: GB-0103424.8
; PRIOR FILING DATE: 2001-02-12
; NUMBER OF SEQ ID NOS: 9218
; SOFTWARE: SeqWin99, version 1.04
; SEQ ID NO 8726
; LENGTH: 459
; TYPE: PRT
; ORGANISM: Neisseria gonorrhoeae
US-10-467-657-8726

Query Match 40.4%; Score 38; DB 6; Length 459;
Best Local Similarity 50.0%; Pred. No. 48;
Matches 8; Conservative 2; Mismatches 6; Indels 0; Gaps 0;

QY 3 ASQDSKINNLTAAQQAQ 18
Db 30 AGEDGKIGNAARQGA 45

RESULT 7

RESULT 9
US-10-821-234-1382
; Sequence 1382, Application US/10821234
; Publication No. US20050255114A1

RESULT 10
US-11-019-711-59
Sequence 59, Application US/11019711
Publication No. US20060009634A1
GENERAL INFORMATION:
APPLICANT: Kekuda, Ramesh
APPLICANT: Alsbrook II, John P
APPLICANT: Tchernev, Velizar T
APPLICANT: Liu, Xiaohong
APPLICANT: Spytek, Kimberly A
APPLICANT: patturajan, Meera
APPLICANT: Grosse, William M
APPLICANT: Lepley, Denise M
APPLICANT: Burgess, Catherine E
APPLICANT: Vernet, Corine A.M.
APPLICANT: Li, Li
APPLICANT: Gorman, Linda
APPLICANT: Edinger, Shlomit R
APPLICANT: Sciore, Paul
APPLICANT: Ellerman, Karen
APPLICANT: Malyankar, Uriel M
APPLICANT: Rothenberg, Mark
APPLICANT: Stone, David J
APPLICANT: Boldog, Ferenc L
APPLICANT: Guo, Xiaojia
APPLICANT: Shenoy, Suresh G
APPLICANT: Anderson, David W
APPLICANT: Padigaru, Muralidhara
APPLICANT: Taupier Jr, Raymond J
APPLICANT: Miller, Charles E
APPLICANT: Eisen, Andrew J
FILE OF INVENTION: Proteins and Nucleic Acids Encoding Same
FILE REFERENCE: 21402-235
CURRENT APPLICATION NUMBER: US/11/019,711
CURRENT FILING DATE: 2004-12-21
PRIOR APPLICATION NUMBER: US/10/037,417
PRIOR FILING DATE: 2002-09-20
PRIOR APPLICATION NUMBER: 60/260,018
PRIOR FILING DATE: 2001-01-05
PRIOR APPLICATION NUMBER: 60/260,360
PRIOR FILING DATE: 2001-01-08
PRIOR APPLICATION NUMBER: 60/272,411
PRIOR FILING DATE: 2001-02-28
PRIOR APPLICATION NUMBER: 60/272,817
PRIOR FILING DATE: 2001-03-02

```

; CURRENT FILING DATE: 2004-02-04
; NUMBER OF SEQ ID NOS: 48640
; SOFTWARE: Patentin version 3.2
; SEQ ID NO 47
; LENGTH: 1050
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-770-726-47

Query Match      39.4%; Score 37; DB 6; Length 1050;
Best Local Similarity 40.0%; Pred. No. 1.8e+02;
Matches      8; Conservative      2; Mismatches 10; Indels      0; Gaps      0;

QY      1 QIASQDSKINNLTAQQAQ 20
      ||||: ||: ||: |
Db      481 QIASQSKIPGMTLSSVCQ 500

RESULT 13
US-10-957-887B-174
; Sequence 174, Application US/10957887B
; Publication No. US20050272677A1
; GENERAL INFORMATION:
; APPLICANT: Friesen, Robert H. E.
; APPLICANT: Leenhouts, Cornelius J.
; APPLICANT: Hektor, Harm
; APPLICANT: van Esch, Johannes H.
; APPLICANT: Heeres, Andre
; TITLE OF INVENTION: DELIVERY OF A SUBSTANCE TO A PRE-DETERMINED SITE
; FILE REFERENCE: 2183-6668US
; CURRENT APPLICATION NUMBER: US/10/957,887B
; CURRENT FILING DATE: 2004-10-04
; PRIOR APPLICATION NUMBER: PCT/NL/00256
; PRIOR FILING DATE: 2003-04-04
; NUMBER OF SEQ ID NOS: 309
; SOFTWARE: Patentin version 3.2
; SEQ ID NO 174
; LENGTH: 45
; TYPE: PRT
; ORGANISM: Escherichia coli O6
US-10-957-887B-174

Query Match      38.3%; Score 36; DB 6; Length 45;
Best Local Similarity 41.2%; Pred. No. 7.1;
Matches      7; Conservative      4; Mismatches 6; Indels      0; Gaps      0;

QY      1 QIASQDSKINNLTAQQQ 17
      ||: ||: ||: ||
Db      25 QLAARDKELRNLIKIGQQ 41

RESULT 14
US-11-145-631-16
; Sequence 16, Application US/11145631
; Publication No. US20060003409A1
; GENERAL INFORMATION:
; APPLICANT: Bertin, John
; TITLE OF INVENTION: NOVEL MOLECULES OF THE CARD-RELATED PROTEIN FAMILY AND
; FILE REFERENCE: 07334-329001
; CURRENT APPLICATION NUMBER: US/11/145,631
; CURRENT FILING DATE: 2005-06-06
; PRIOR APPLICATION NUMBER: US/09/841,739
; PRIOR FILING DATE: 2001-08-29
; PRIOR APPLICATION NUMBER: US 09/697,089
; PRIOR FILING DATE: 2000-10-26
; PRIOR APPLICATION NUMBER: US 60/161,822
; PRIOR FILING DATE: 1999-10-27
; NUMBER OF SEQ ID NOS: 16
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 16
; LENGTH: 89
; TYPE: PRT
; ORGANISM: Homo sapiens

```

US-11-145-631-16

```
Query Match      38.3%; Score 36; DB 7; Length 89;
Best Local Similarity 43.8%; Pred. NO. 16;
Matches 7; Conservative 3; Mismatches 6; Indels 0; Gaps 0;
```

Qy	4	SQDSKINNLTAAQQA	19
		: : : : :	
Db	38	SEEEKVRNEPTQQORA	53

RESULT 15

```

US-11-074-176-136
; Sequence 136, Application US/11074176
; Publication No. US20050250135A1
; GENERAL INFORMATION:
; APPLICANT: Klaenhammer, Todd R.
; APPLICANT: Russell, William M.
; APPLICANT: Altermann, Eric
; APPLICANT: McAuliffe, Olivia
; APPLICANT: Peril, Andrea Azcarate
; TITLE OF INVENTION: Nucleic Acid Sequences Encoding
; TITLE OF INVENTION: Stress-Related Proteins and Uses Therefore
; FILE REFERENCE: 5051-694
; CURRENT APPLICATION NUMBER: US/11/074,176
; CURRENT FILING DATE: 2005-03-07
; PRIOR APPLICATION NUMBER: 60/551,161
; PRIOR FILING DATE: 2004-03-08
; NUMBER OF SEQ ID NOS: 381
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 136
; LENGTH: 228
; TYPE: PRT
; ORGANISM: Lactobacillus acidophilus
US-11-074-176-136

```

Query Match 38.3%; Score 36; DB 7; Length 228;
Best Local Similarity 40.0%; Pred. No. 46;
Matches 8; Conservative 4; Mismatches 8; Indels 0; Gaps 0;

```

QY      1 QIASQDSKINNLTAQQAAQ 20
        | : | | | :
Db     202 QVQTQGEHNKKAEEQAAK 221

```

Search completed: January 25, 2006, 20:11:08
Job time : 5.5 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2006 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: January 25, 2006, 18:48:24 ; Search time 68.2 Seconds
(without alignments)
128.850 Million cell updates/sec

Title: US-10-797-821-7
Perfect score: 96
Sequence: 1 QDSKINLTAQQAAQAQVN 20

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 2443163 seqs, 439378781 residues

Total number of hits satisfying chosen parameters: 2443163

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : A_Geneseq_21.*
1: Geneseqp1980s:*
2: Geneseqp1990s:*
3: Geneseqp2000s:*
4: Geneseqp2001s:*
5: Geneseqp2002s:*
6: Geneseqp2003as:*
7: Geneseqp2003bs:*
8: Geneseqp2004s:*
9: Geneseqp2005s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	96	100.0	20	7	AdD93627 Streptoco
2	96	100.0	20	9	AdX37250 Streptoco
3	96	100.0	431	7	AdD93649 Streptoco
4	96	100.0	431	7	AdD93650 Streptoco
5	96	100.0	431	7	AdD93653 Streptoco
6	96	100.0	431	9	AdX37272 Streptoco
7	96	100.0	431	9	AdX37273 Streptoco
8	96	100.0	431	9	AdX37276 Streptoco
9	96	100.0	431	9	Aeb91500 Microbial
10	96	100.0	432	7	AdD93651 Streptoco
11	96	100.0	432	7	AdD93652 Streptoco
12	96	100.0	432	9	AdX37274 Streptoco
13	96	100.0	432	9	AdX37275 Streptoco
14	77	80.2	20	7	AdD93626 Streptoco
15	77	80.2	20	9	AdX37249 Streptoco
16	73	76.0	210	2	Aay22579 Bacttrial
17	73	76.0	392	6	Abu02747 S. pneumo
18	73	76.0	392	8	Adk47859 Streptoco
19	73	76.0	392	8	Adt50227 S.pneumon
20	73	76.0	392	8	Adt50226 S.pneumon
21	73	76.0	392	8	Adt50165 S.pneumon
22	73	76.0	399	8	AdR94595 Novel S.
23	73	76.0	399	9	Aea58465 Streptoco
24	61	63.5	398	5	Abp25919 Streptoco

ALIGNMENTS

RESULT 1

AD	ADD93627	standard; peptide; 20 AA.	
XX	XX	ADD93627;	
XX	XX	29-JAN-2004 (first entry)	
XX	XX	Streptococcus mutans glucan binding protein-B peptide fragment.	
XX	XX	Glucan binding protein-B; GbpB; vaccine; anticaries; epitope; immunogen.	
XX	XX	Streptococcus mutans.	
XX	XX	WO2003075845-A2.	
XX	XX	18-SEP-2003.	
XX	XX	07-MAR-2003; 2003WO-US006962.	
XX	XX	07-MAR-2002; 2002US-0363209P.	
XX	XX	08-AUG-2002; 2002US-0402483P.	
XX	XX	(FORS-) FORSYTH INST.	
XX	XX	Smith DJ, Taubman MA;	
XX	XX	WPI; 2003-845091/78.	
XX	XX	Composition useful as vaccines for dental caries comprises a fragment of a glucan binding protein-B binding to a major histocompatibility complex class II protein.	
XX	XX	Claim 11; Page 10; 49pp; English.	
XX	XX	The present sequence is that of a peptide comprising amino acid residues 37-56 of the glucan binding protein-B (GbpB) of Streptococcus mutans strain SK32 ADD93649. The peptide binds to a major histocompatibility complex (MHC) class II protein. It was identified as a potential B cell epitope using a matrix-based algorithm for epitope prediction, which was used to search the primary amino acid sequence of GbpB for known MHC class II binding motifs. The peptide can be used in immunogenic compositions and subunit vaccines for dental caries. These comprise MHC class II protein-binding GbpB peptides covalently linked with peptide subunits (preferably from the catalytic domain) of a glucosyltransferase. The compositions are used in a claimed method of eliciting production of an antibody in a mammal. Dieptopic or multieptopic polypeptides can be	

Adr3884 S. pyogen
Aeb91542 Microbial
Abb57775 Drosophil
Abr40385 Fruitfly
Adc96412 Drosophil
Adc06779 Rat fly
Abp29684 Streptoco
Adu69524 S agalact
Adv88392 Streptoco
Adv81808 Streptoco
Adv79645 Streptoco
Adv16553 E. faecal
Adh88105 Enterococ
Abg75182 M licheni
Abu43868 Protein e
Aab29586 Rat potas
Aau00216 Rat potas
Abb65663 Drosophil
Abg99869 S. cinnam
Abu29734 Protein e
Adc95468 E. faeciu

25 61 63.5 398 8 ADR83884
26 61 63.5 398 9 AEB91542
27 54 56.2 325 4 ABB57775
28 54 56.2 325 6 ABR40385
29 54 56.2 325 8 ADS96412
30 54 56.2 478 7 ADC06779
31 52 54.2 447 5 ADP29684
32 52 54.2 447 8 ADU69524
33 52 54.2 447 8 ADV88392
34 52 54.2 447 8 ADV81808
35 52 54.2 447 8 ADV79645
36 52 54.2 482 9 ADV16553
37 52 54.2 497 7 ADH88105
38 52 54.2 1050 7 ABG75182
39 50.5 52.6 1188 6 ABU43868
40 48 50.0 123 3 AAB29586
41 48 50.0 123 4 AAU00216
42 48 50.0 618 4 ABB65663
43 47 49.0 2238 4 ABG99869
44 46 47.9 524 6 ABU29734
45 46 47.9 525 7 ADC95468

CC prepared synthetically or by recombinant DNA technology. Antibodies
 CC raised against MHC class II binding fragments of GbpB can be used in
 CC passive immunisation.

SQ Sequence 20 AA;

Query Match 100.0%; Score 96; DB 7; Length 20;
 Best Local Similarity 100.0%; Pred. No. 7.8e-08;
 Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 QDSKINNLTAAQQAAQAQVN 20
 |||||
 Db 1 QDSKINNLTAAQQAAQAQVN 20

RESULT 2

ADX37250
 ID ADX37250 standard; peptide; 20 AA.

XX AC ADX37250;

XX DT 21-APR-2005 (first entry)

XX DE Streptococcus mutant glucan binding protein B peptide #7.

XX KW immunogenicity; immune stimulation; glucan binding protein-B;
 KW microparticle; major histocompatibility complex; tooth disease.

XX OS Streptococcus mutans.

XX PN US2005031633-A1.

XX PD 10-FEB-2005.

XX PF 09-MAR-2004; 2004US-00797821.

XX PR 13-APR-1998; 98US-0081550P.

XX PR 08-JAN-1999; 99US-0115142P.

XX PR 12-APR-1999; 99US-00290049.

XX PR 07-MAR-2002; 2002US-0363209P.

XX PR 08-AUG-2002; 2002US-0402483P.

XX PR 07-MAR-2003; 2003US-00383930.

XX PA (SMIT/) SMITH D J.

XX PA (TAUB/) TAUBMAN M A.

XX PI Smith DJ, Taubman MA;

XX PI WPI; 2005-151644/16.

XX DR Claim 4; SEQ ID NO 7; 73pp; English.

XX PT New composition comprising a fragment of a glucan binding protein-B
 PT (GbpB) that binds to MHC class II protein, and a biocompatible
 PT microparticle, useful for producing an antibody (claimed) for immunizing
 PT mammals against dental caries.

XX PS Claim 4; SEQ ID NO 7; 73pp; English.

XX CC The invention relates to a composition comprising a fragment of a glucan
 CC binding protein-B (GbpB) and a biocompatible microparticle, where the
 CC fragment binds to a major histocompatibility complex (MHC) class II
 CC protein. The composition is useful for producing an antibody for
 CC immunizing mammals against dental caries. This sequence corresponds to a
 CC Streptococcus mutans GbpB-derived peptide of the invention.

XX SQ Sequence 20 AA;

Query Match 100.0%; Score 96; DB 9; Length 20;
 Best Local Similarity 100.0%; Pred. No. 7.8e-08;
 Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 QDSKINNLTAAQQAAQAQVN 20
 |||||

Db 1 QDSKINNLTAAQQAAQAQVN 20

RESULT 3
 ADD93649
 ID ADD93649 standard; protein; 431 AA.
 XX AC ADD93649;
 XX DT 29-JAN-2004 (first entry)
 XX DE Streptococcus mutans glucan binding protein-B.
 XX KW Glucan binding protein-B; GbpB; vaccine; anticaries; epitope; immunogen.
 XX OS Streptococcus mutans.
 XX FH Key Location/Qualifiers
 FT Region 6..25
 FT /note= "HLA-binding peptide"
 FT Region 16..35
 FT /note= "HLA-binding peptide"
 FT Region 33..52
 FT /note= "HLA-binding peptide"
 FT Region 37..56
 FT /note= "HLA-binding peptide"
 FT Region 48..67
 FT /note= "HLA-binding peptide"
 FT Region 52..71
 FT /note= "HLA-binding peptide"
 FT Region 88..107
 FT /note= "HLA-binding peptide"
 FT Region 113..132
 FT /note= "HLA-binding peptide"
 FT Region 117..136
 FT /note= "HLA-binding peptide"
 FT Region 137..156
 FT /note= "HLA-binding peptide"
 FT Region 174..193
 FT /note= "HLA-binding peptide"
 FT Region 194..213
 FT /note= "HLA-binding peptide"
 FT Region 214..233
 FT /note= "HLA-binding peptide"
 FT Region 248..267
 FT /note= "HLA-binding peptide"
 FT Region 289..308
 FT /note= "HLA-binding peptide"
 FT Region 306..325
 FT /note= "HLA-binding peptide"
 FT Region 311..330
 FT /note= "HLA-binding peptide"
 FT Region 349..368
 FT /note= "HLA-binding peptide"
 FT Region 365..384
 FT /note= "HLA-binding peptide"
 FT Region 383..402
 FT /note= "HLA-binding peptide"
 FT Region 403..422
 FT /note= "HLA-binding peptide"
 XX WO2003075845-A2.
 XX PD 18-SEP-2003.
 XX PF 07-MAR-2003; 2003WO-US006962.
 XX PR 07-MAR-2002; 2002US-0363209P.
 XX PR 08-AUG-2002; 2002US-0402483P.
 XX PA (FORS-) FORSYTH INST.
 XX PI Smith DJ, Taubman MA;
 XX

CC carries. The compositions comprise major histocompatibility complex (MHC) class II protein-binding GbpB peptides, especially HLA-binding peptides, covalently linked with peptide subunits of a glucosyltransferase. The compositions are used in a claimed method of eliciting production of an antibody in a mammal. Dieptopic or multi-epitopic polypeptides can be prepared synthetically or by recombinant DNA technology. Antibodies raised against MHC class II binding fragments of GbpB can be used in passive immunisation.

XX SQ Sequence 431 AA;

Query Match 100.0%; Score 96; DB 7; Length 431;
Best Local Similarity 100.0%; Pred. No. 2.4e-06;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0

QY 1 QDSKNNLTAQQQAAQAVN 20
|||||

DB 37 QDSKNNLTAQQQAAQAVN 56
|||||

RESULT 5

ADD93653

ID ADD93653 standard; protein; 431 AA.

XX

AC ADD93653;

XX

XX 29-JAN-2004 (first entry)

XX Streptococcus mutans glucan binding protein-B.

XX Glucan binding protein-B; GbpB; vaccine; anticaries; epitope; immunogen.

XX Streptococcus mutans.

OS

XX WO2003075845-A2.

PN

XX 18-SEP-2003.

PD

XX

XX 07-MAR-2003; 2003WO-US006962.

XX

XX 07-MAR-2002; 2002US-0363209P.

PR

XX 08-AUG-2002; 2002US-0402483P.

PR

XX (FORS-) FORSYTH INST.

XX

XX Smith DJ, Taubman MA;

PI

XX WPI; 2003-845091/78.

DR

XX GENBANK; AY046414.

DR

XX Composition useful as vaccines for dental caries comprises a fragment of a glucan binding protein-B binding to a major histocompatibility complex class II protein.

XX

XX Claim 5; Page 8-9; 49pp; English.

XX

XX The present sequence is the protein sequence of the glucan binding protein-B (GbpB) of Streptococcus mutans strain 55M3. The sequence includes a number of human leucocyte antigen (HLA)-binding epitopes. The invention provides immunogenic compositions and vaccines for dental caries. The compositions comprise major histocompatibility complex (MHC) class II protein-binding GbpB peptides, especially HLA-binding peptides, covalently linked with peptide subunits of a glucosyltransferase. The compositions are used in a claimed method of eliciting production of an antibody in a mammal. Dieptopic or multi-epitopic polypeptides can be prepared synthetically or by recombinant DNA technology. Antibodies raised against MHC class II binding fragments of GbpB can be used in passive immunisation.

XX SQ Sequence 431 AA;

Query Match 100.0%; Score 96; DB 7; Length 431;
Best Local Similarity 100.0%; Pred. No. 2.4e-06;

PF 09-MAR-2004; 2004US-00797821.
 XX 13-APR-1998; 98US-0081550P.
 PR 08-JAN-1999; 99US-0115142P.
 PR 12-APR-1999; 99US-00290049.
 PR 07-MAR-2002; 2002US-0363209P.
 PR 08-AUG-2002; 2002US-0402483P.
 PR 07-MAR-2003; 2003US-00383930.
 XX (SMIT/) SMITH D J.
 PA (TAUB/) TAUBMAN M A.
 XX Smith DJ, Taubman MA;
 PI WPI; 2005-151644/16.
 DR New composition comprising a fragment of a glucan binding protein-B
 XX (GbpB) that binds to MHC class II protein, and a biocompatible
 PT microparticle, useful for producing an antibody (claimed) for immunizing
 PT mammals against dental caries.
 XX Claim 3; SEQ ID NO 33; 73pp; English.
 PS The invention relates to a composition comprising a fragment of a glucan
 XX binding protein-B (GbpB) and a biocompatible microparticle, where the
 CC fragment binds to a major histocompatibility complex (MHC) class II
 CC protein. The composition is useful for producing an antibody for
 CC immunizing mammals against dental caries. This sequence corresponds to a
 CC Streptococcus mutans GbpB protein of the invention.
 XX Sequence 431 AA;
 SQ Query Match 100.0%; Score 96; DB 9; Length 431;
 Best Local Similarity 100.0%; Pred. No. 2.4e-06;
 Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 QDSKINLTAQQAQAQVN 20
 DB ||||||||||||||||||
 37 QDSKINLTAQQAQAQVN 56
 RESULT 9
 ID AEB91500 standard; protein; 431 AA.
 XX AEB91500;
 AC 20-OCT-2005 (first entry)
 DT Microbial pathogen adhesin protein sequence, SEQ ID NO:210.
 XX algorithm; adhesin; pharmaceutical; vaccine; drug screening;
 KW bordetella pertussis infection; antibacterial; pneumonia;
 KW antiinflammatory; respiratory-gen.; gastric ulcer; antilucer;
 KW gastrointestinal-gen.; urinary tract infection; antimicrobial; uropathic.
 XX Streptococcus mutans.
 OS WO2005076010-A2.
 PN 18-AUG-2005.
 PD 07-FEB-2005; 2005WO-IN000037.
 XX 06-FEB-2004; 2004IN-DE000173.
 PR 20-JUL-2004; 2004US-0589227P.
 XX (COUL) COUNCIL SCI & IND RES SOUTH AFRICA.
 PA Sachdeva G, Kumar K, Jain P, Brahmachari SK, Ramachandran S;
 PI WPI; 2005-597835/61.
 XX

PT Computational method for identifying adhesin and adhesin like molecules,
 PT comprises computing sequence-based attributes of protein sequences using
 XX neural network software and training an artificial neural network.
 PS Claim 16; SEQ ID NO 210; 402pp; English.
 XX The present invention relates to a computational method (M1) for
 CC identifying adhesin and adhesin-like proteins, by computing the sequence-
 CC based attributes of protein sequences using five attribute modules of a
 CC neural network software, training an artificial neural network (ANN) for
 CC each of the computed five attributes, and identifying the adhesin and
 CC adhesin-like proteins having probability of being an adhesin (Pad) as
 CC equal or greater than 0.51. Also claimed is a set of 274 annotated genes
 CC encoding adhesin and adhesin-like proteins, having 274 fully defined 162-
 CC 1151 base pair (SEQ ID NO: 385-658) sequences; a set of 105 hypothetical
 CC genes encoding adhesin and adhesin-like proteins, having 105 fully
 CC defined 306-15876 base pairs (SEQ ID NO: 659-763); a set of 279 annotated
 CC adhesin and adhesin-like proteins, having 279 fully defined 53-3716 base
 CC pair (SEQ ID NO: 1-279); a set of 105 hypothetical adhesin and adhesin-
 CC like proteins, having 105 fully defined 106-5291 base pair (SEQ ID NO:
 CC 280-384) sequences; and a fully connected multilayer feed forward ANN (1)
 CC based on (M1). (M1) is useful for identifying adhesin and adhesin-like
 CC proteins, of therapeutic potential, and identifying and short-listing
 CC to eliminate diseases caused by various pathogenic organisms. (M1) is
 CC useful for identifying putative adhesins that are important in drug
 CC discovery and preventing therapeutics for whooping cough, pneumonia,
 CC gastric ulcer and urinary tract infections. (M1) identifies adhesins from
 CC distantly related organisms, and from bacteria belonging to a wide
 CC phylogenetic spectrum. (M1) is capable of predicting adhesive nature of
 CC unique proteins. The present sequence is a microbial pathogen adhesin
 CC protein sequence.
 XX Sequence 431 AA;
 SQ Query Match 100.0%; Score 96; DB 9; Length 431;
 Best Local Similarity 100.0%; Pred. No. 2.4e-06;
 Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 QDSKINLTAQQAQAQVN 20
 DB ||||||||||||||||||
 37 QDSKINLTAQQAQAQVN 56
 RESULT 10
 ID ADD93651 standard; protein; 432 AA.
 XX ADD93651;
 AC 29-JAN-2004 (first entry)
 DT Streptococcus mutans glucan binding protein-B.
 XX Glucan binding protein-B; GbpB; vaccine; anticaries; epitope; immunogen.
 KW Streptococcus mutans.
 OS WO2003075845-A2.
 PN 18-SEP-2003.
 PD 07-MAR-2003; 2003WO-US006962.
 XX 07-MAR-2002; 2002US-0363209P.
 PR 08-AUG-2002; 2002US-0402483P.
 XX (FORS-) FORSYTH INST.
 PA Smith DJ, Taubman MA;
 PI WPI; 2003-845091/78.
 XX GENBANK; AY046412.
 DR

XX Composition useful as vaccines for dental caries comprises a fragment of
PT a glucan binding protein-B binding to a major histocompatibility complex
PT class II protein.
XX
XX Claim 5; Page 8; 49pp; English.
XX
CC The present sequence is the protein sequence of the glucan binding
CC protein-B (GbpB) of Streptococcus mutans strain 15P2. The sequence
CC includes a number of human leucocyte antigen (HLA)-binding epitopes. The
CC invention provides immunogenic compositions and vaccines for dental
CC caries. The compositions comprise major histocompatibility complex (MHC)
CC class II protein-binding GbpB peptides, especially HLA-binding peptides,
CC covalently linked with peptide subunits of a glucosyltransferase. The
CC compositions are used in a claimed method of eliciting production of an
CC antibody in a mammal. Dieptopic or multiepitopic polypeptides can be
CC prepared synthetically or by recombinant DNA technology. Antibodies
CC raised against MHC class II binding fragments of GbpB can be used in
CC passive immunisation.
XX
XX Sequence 432 AA;
SQ
Query Match 100.0%; Score 96; DB 7; Length 432;
Best Local Similarity 100.0%; Pred. No. 2.4e-06;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 QDSKINLTAQQAAQAQVN 20
Db 37 QDSKINLTAQQAAQAQVN 56
RESULT 11
ADD93652
ID ADD93652 standard; protein; 432 AA.
XX
XX AC ADD93652;
XX
XX DT 29-JAN-2004 (first entry)
XX
XX DE Streptococcus mutans glucan binding protein-B.
XX
XX KW Glucan binding protein-B; GbpB; vaccine; anticaries; epitope; immunogen.
XX
XX OS Streptococcus mutans.
XX
XX PN WO2003075845-A2.
XX
XX PD 18-SEP-2003.
XX
XX PF 07-MAR-2003; 2003WO-US006962.
XX
XX PR 07-MAR-2002; 2002US-0363209P.
XX
XX PR 08-AUG-2002; 2002US-0402483P.
XX
XX PA (FORS-) FORSYTH INST.
XX
XX PI Smith DJ, Taubman MA;
XX
XX PS WPI; 2003-845091/78.
XX
XX DR GENBANK; AY046413.
XX
XX PT Composition useful as vaccines for dental caries comprises a fragment of
PT a glucan binding protein-B binding to a major histocompatibility complex
PT class II protein.
XX
XX PS Claim 5; Page 8; 49pp; English.
XX
XX CC The present sequence is the protein sequence of the glucan binding
CC protein-B (GbpB) of Streptococcus mutans strain 3SM1. The sequence
CC includes a number of human leucocyte antigen (HLA)-binding epitopes. The
CC invention provides immunogenic compositions and vaccines for dental
CC caries. The compositions comprise major histocompatibility complex (MHC)
CC class II protein-binding GbpB peptides, especially HLA-binding peptides,
CC covalently linked with peptide subunits of a glucosyltransferase. The
CC compositions are used in a claimed method of eliciting production of an
CC antibody in a mammal. Dieptopic or multiepitopic polypeptides can be
CC prepared synthetically or by recombinant DNA technology. Antibodies
CC raised against MHC class II binding fragments of GbpB can be used in
CC passive immunisation.
XX
XX Sequence 432 AA;
SQ
Query Match 100.0%; Score 96; DB 7; Length 432;
Best Local Similarity 100.0%; Pred. No. 2.4e-06;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 QDSKINLTAQQAAQAQVN 20
Db 37 QDSKINLTAQQAAQAQVN 56
RESULT 12
ADD937274
ID ADD937274 standard; protein; 432 AA.
XX
XX AC ADD937274;
XX
XX DT 21-APR-2005 (first entry)
XX
XX DE Streptococcus mutans glucan binding protein B variant #3.
XX
XX KW immunogenicity; immune stimulation; glucan binding protein-B;
KW microparticle; major histocompatibility complex; tooth disease.
XX
XX OS Streptococcus mutans.
XX
XX PN US2005031633-A1.
XX
XX PD 10-FEB-2005.
XX
XX PF 09-MAR-2004; 2004US-00797821.
XX
XX PR 13-APR-1998; 98US-0081550P.
XX
XX PR 08-JAN-1999; 99US-0115142P.
XX
XX PR 12-APR-1999; 99US-0029004P.
XX
XX PR 07-MAR-2002; 2002US-0363209P.
XX
XX PR 08-AUG-2002; 2002US-0402483P.
XX
XX PR 07-MAR-2003; 2003US-00363930.
XX
XX PA (SMIT/) SMITH D J.
XX
XX PA (TAUB/) TAUBMAN M A.
XX
XX PI Smith DJ, Taubman MA;
XX
XX DR WPI; 2005-151644/16.
XX
XX PT New composition comprising a fragment of a glucan binding protein-B
PT (GbpB) that binds to MHC class II protein, and a biocompatible
PT microparticle, useful for producing an antibody (claimed) for immunizing
PT mammals against dental caries.
XX
XX PS Claim 3; SEQ ID NO 31; 73pp; English.
XX
XX CC The invention relates to a composition comprising a fragment of a glucan
CC binding protein-B (GbpB) and a biocompatible microparticle, where the
CC fragment binds to a major histocompatibility complex (MHC) class II
CC protein. The composition is useful for producing an antibody for
CC immunizing mammals against dental caries. This sequence corresponds to a
CC Streptococcus mutans GbpB protein of the invention.
XX
XX SQ Sequence 432 AA;
Query Match 100.0%; Score 96; DB 9; Length 432;
Best Local Similarity 100.0%; Pred. No. 2.4e-06;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 QDSKINNLTAQQQAQAQVN 20
 DB 37 QDSKINNLTAQQQAQAQVN 56

RESULT 13
 ADX37275
 ID ADX37275 standard; protein; 432 AA.
 AC ADX37275;
 XX
 DT 21-APR-2005 (first entry)
 XX
 DE Streptococcus mutant glucan binding protein B variant #4.
 XX
 KW immunogenicity; immune stimulation; glucan binding protein-B;
 KW microparticle; major histocompatibility complex; tooth disease.
 XX
 OS Streptococcus mutans.
 XX
 PN US2005031633-A1.
 XX
 PD 10-FEB-2005.
 XX
 PF 09-MAR-2004; 2004US-00797821.
 XX
 PR 13-APR-1998; 98US-0081550P.
 PR 08-JAN-1999; 99US-0115142P.
 PR 12-APR-1999; 99US-00290049.
 PR 07-MAR-2002; 2002US-0363209P.
 PR 08-AUG-2002; 2002US-0402483P.
 PR 07-MAR-2003; 2003US-00363930.
 XX
 PA (SMIT/) SMITH D J.
 PA (TAUB/) TAUBMAN M A.
 PI Smith DJ, Taubman MA;
 PI Smith DJ, Taubman MA;
 DR WPI; 2005-151644/16.
 XX
 PT New composition comprising a fragment of a glucan binding protein-B
 PT (GbpB) that binds to MHC class II protein, and a biocompatible
 PT microparticle, useful for producing an antibody (claimed) for immunizing
 PT mammals against dental caries.
 XX
 PS Claim 3; SEQ ID NO 32; 73pp; English.
 XX
 CC The invention relates to a composition comprising a fragment of a glucan
 CC binding protein-B (GbpB) and a biocompatible microparticle, where the
 CC fragment binds to a major histocompatibility complex (MHC) class II
 CC protein. The composition is useful for producing an antibody for
 CC immunizing mammals against dental caries. This sequence corresponds to a
 CC Streptococcus mutans GbpB protein of the invention.
 XX
 SQ Sequence 432 AA;

Query Match 100.0%; Score 96; DB 9; Length 432;
 Best Local Similarity 100.0%; Pred. No. 2.4e-06;
 Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 QDSKINNLTAQQQAQAQVN 20
 DB 37 QDSKINNLTAQQQAQAQVN 56

RESULT 14
 ADD93626
 ID ADD93626 standard; peptide; 20 AA.
 XX
 AC ADD93626;
 XX
 DT 29-JAN-2004 (first entry)
 XX

DE Streptococcus mutans glucan binding protein-B peptide fragment.
 XX
 KW Glucan binding protein-B; GbpB; vaccine; anticaries; epitope; immunogen.
 XX
 OS Streptococcus mutans.
 XX
 PN WO2003075845-A2.
 XX
 PD 18-SEP-2003.
 XX
 PF 07-MAR-2003; 2003WO-US006962.
 XX
 PR 07-MAR-2002; 2002US-0363209P.
 PR 08-AUG-2002; 2002US-0402483P.
 XX
 PA (FORS-) FORSYTH INST.
 PI Smith DJ, Taubman MA;
 PI Smith DJ, Taubman MA;
 DR WPI; 2003-845091/78.
 XX
 PT Composition useful as vaccines for dental caries comprises a fragment of
 PT a glucan binding protein-B binding to a major histocompatibility complex
 PT class II protein.
 XX
 PS Claim 11; Page 10; 49pp; English.
 XX
 CC The present sequence is that of a peptide comprising amino acid residues
 CC 33-52 of the glucan binding protein-B (GbpB) of Streptococcus mutans
 CC strain SK32 ADD93649. The peptide binds to a major histocompatibility
 CC complex (MHC) class II protein. It was identified as a potential B cell
 CC epitope using a matrix-based algorithm for epitope prediction, which was
 CC used to search the primary amino acid sequence of GbpB for known MHC
 CC class II binding motifs. The peptide can be used in immunogenic
 CC compositions and subunit vaccines for dental caries. These comprise MHC
 CC class II protein-binding GbpB peptides covalently linked with peptide
 CC subunits (preferably from the catalytic domain) of a glucosyltransferase.
 CC The compositions are used in a claimed method of eliciting production of
 CC an antibody in a mammal. Dieptopic or multiepitopic polypeptides can be
 CC prepared synthetically or by recombinant DNA technology. Antibodies
 CC raised against MHC class II binding fragments of GbpB can be used in
 CC passive immunisation.
 XX
 SQ Sequence 20 AA;

Query Match 80.2%; Score 77; DB 7; Length 20;
 Best Local Similarity 100.0%; Pred. No. 6.9e-05;
 Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 QDSKINNLTAQQQAQAQ 16
 DB 5 QDSKINNLTAQQQAQAQ 20

RESULT 15
 ADX37249
 ID ADX37249 standard; peptide; 20 AA.
 XX
 AC ADX37249;
 XX
 DT 21-APR-2005 (first entry)
 XX
 DE Streptococcus mutant glucan binding protein B peptide #6.
 XX
 KW immunogenicity; immune stimulation; glucan binding protein-B;
 KW microparticle; major histocompatibility complex; tooth disease.
 XX
 OS Streptococcus mutans.
 XX
 PN US2005031633-A1.
 XX
 PD 10-FEB-2005.
 XX

PF 09-MAR-2004; 2004US-00797821.
XX
PR 13-APR-1998; 98US-0081550P.
PR 08-JAN-1999; 99US-0115142P.
PR 12-APR-1999; 99US-00290049.
PR 07-MAR-2002; 2002US-0363209P.
PR 08-AUG-2002; 2002US-0402483P.
PR 07-MAR-2003; 2003US-00383930.
XX
PA (SMIT/) SMITH D J.
PA (TAUB/) TAUBMAN M A.
XX
PI Smith DJ, Taubman MA;
XX
DR WPI; 2005-151644/16.
XX
PT New composition comprising a fragment of a glucan binding protein-B
PT (GbpB) that binds to MHC class II protein, and a biocompatible
PT microparticle, useful for producing an antibody (claimed) for immunizing
PT mammals against dental caries.
XX
PS Claim 4; SEQ ID NO 6; 73pp; English.
XX
CC The invention relates to a composition comprising a fragment of a glucan
CC binding protein-B (GbpB) and a biocompatible microparticle, where the
CC fragment binds to a major histocompatibility complex (MHC) class II
CC protein. The composition is useful for producing an antibody for
CC immunizing mammals against dental caries. This sequence corresponds to a
CC Streptococcus mutans GbpB-derived peptide of the invention.
XX
SQ Sequence 20 AA;

Query Match 80.2%; Score 77; DB 9; Length 20;
Best Local Similarity 100.0%; Pred. No. 6.9e-05;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 QDSKINNLTAQQQAAQ 16
| | | | | | | | | | | | | | | |
Db 5 QDSKINNLTAQQQAAQ 20

Search completed: January 25, 2006, 19:01:44
Job time : 69.2 secs

Query Match 76.0%; Score 73; DB 2; Length 392;
Best Local Similarity 75.0%; Pred. No. 0.00058;
Matches 15; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

QY 1 QDSKNNLTAQQQAQAQVN 20
||:|||||:|||||:
Db 37 QDNKISNLTAQQQAQAQKQVD 56

RESULT 3
S05542
hypothetical protein, 54K - Enterococcus faecium
C:Species: Enterococcus faecium
C:Date: 07-Sep-1990 #sequence_revision 07-Sep-1990 #text_change 09-Jul-2004
C:Accession: S05542
R:Fuerst, P.; Moesch, H.U.; Solioz, M.
Nucleic Acids Res. 17, 6724, 1989
A:Title: A protein of unusual composition from Enterococcus faecium.
A:Reference number: S05542; MUID:89385998; PMID:2780297
A:Accession: S05542
A:Molecule type: DNA
A:Residues: 1-507 <FUS>
A:Cross-references: UNIPROT:P13692; UNIPARC:UPI000016F6FC; GB:X16421; EMBL:M26048; NID:9
A>Note: the authors translated the codon CGT for residues 221 and 223 as Lys

Query Match 50.0%; Score 48; DB 2; Length 507;
Best Local Similarity 52.6%; Pred. No. 7.5;
Matches 10; Conservative 4; Mismatches 5; Indels 0; Gaps 0;

QY 1 QDSKNNLTAQQQAQAQV 19
|||:||||:||||:
Db 28 QDKKIADLQNQQAQAQSQI 46

RESULT 4
C64189
succinate-CoA ligase (ADP-forming) (EC 6.2.1.5) beta chain - Haemophilus influenzae (str
C:Species: Haemophilus influenzae
C:Date: 18-Aug-1995 #sequence_revision 18-Aug-1995 #text_change 05-Oct-2004
C:Accession: C64189
R:Flaeschmann, R.D.; Adams, M.D.; White, O.; Clayton, R.A.; Kirkness, E.F.; Kerlavage, A
; Gocayne, J.D.; Scott, J.; Shirley, R.; Liu, L.I.; Glodek, A.; Kelley, J.M.; Weidman, J
; D.M.; Brandon, R.C.; Fine, L.D.; Fritchman, J.L.; Fuhrmann, J.L.; Geoghagen, N.S.M.
Science 269, 496-512, 1995
A:Authors: Gnehm, C.L.; McDonald, L.A.; Small, K.V.; Fraser, C.M.; Smith, H.O.; Venter,
A:Title: Whole-genome random sequencing and assembly of Haemophilus influenzae Rd.
A:Reference number: A64000; MUID:95350630; PMID:7542800
A:Accession: C64189
A:Status: nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-389 <TRIGR>
A:Cross-references: UNIPROT:P45101; UNIPARC:UPI000013619E; GB:U32799; GB:L42023; NID:g15
C:Superfamily: succinyl-CoA synthetase, beta subunit
C:Keywords: acid-thiol ligase; coenzyme A

Query Match 47.9%; Score 46; DB 2; Length 389;
Best Local Similarity 57.1%; Pred. No. 12;
Matches 12; Conservative 3; Mismatches 4; Indels 2; Gaps 1;

QY 2 DSKNNLTAQ--QQAQAQVN 20
:|:|||||:|||||:
Db 363 ESDVNILTAQSLQQAALAVN 383

RESULT 5
S36575
EI protein - human papillomavirus type 52
C:Species: human papillomavirus type 52
C:Date: 20-Feb-1995 #sequence_revision 20-Feb-1995 #text_change 09-Jul-2004
C:Accession: S36575
R:Delius, H.; Hofmann, B.
submitted to the EMBL Data Library, August 1993
A:Description: Primer-directed sequencing of human papillomavirus types.

A:Reference number: S36469
A:Accession: S36575
A:Molecule type: DNA
A:Residues: 1-647
A:Cross-references: UNIPROT:P36730; UNIPARC:UPI0000138310; EMBL:X74481; NID:g397038; PFI
C:Superfamily: papillomavirus E1 protein
C:Keywords: early protein; nucleus

Query Match 47.9%; Score 46; DB 2; Length 647;
Best Local Similarity 52.6%; Pred. No. 20;
Matches 10; Conservative 3; Mismatches 6; Indels 0; Gaps 0;

QY 2 DSKNNLTAQQQAQAQVN 20
|||||:|:|||||:
Db 53 DSNINNEQAHEHAARALFN 71

RESULT 6
H81850
probable membrane lipoprotein NMA1581 [imported] - Neisseria meningitidis (strain Z2491
C:Species: Neisseria meningitidis
C:Date: 05-May-2000 #sequence_revision 05-May-2000 #text_change 09-Jul-2004
C:Accession: H81850
R:Paikhill, J.; Achtman, M.; James, K.D.; Bentley, S.D.; Churcher, C.; Klee, S.R.; Morel
; Holroyd, S.; Jagels, K.; Leather, S.; Moule, S.; Mungall, K.; Quail, M.A.; Rajandream,
Nature 404, 502-506, 2000
A:Title: Complete DNA sequence of a serogroup A strain of Neisseria meningitidis Z2491.
A:Reference number: A81775; MUID:20222556; PMID:10761919
A:Accession: H81850
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-184 <PAR>
A:Cross-references: UNIPROT:Q9JTY2; UNIPARC:UPI00000C4C19; GB:AL162756; GB:AL157959; NII
A:Experimental source: serogroup A, strain Z2491
C:Genetics:
C:Gene: NMA1581
C:Superfamily: Neisseria meningitidis probable membrane lipoprotein NMA1581

Query Match 46.9%; Score 45; DB 2; Length 184;
Best Local Similarity 55.6%; Pred. No. 7.7;
Matches 10; Conservative 3; Mismatches 5; Indels 0; Gaps 0;

QY 1 QDSKNNLTAQQQAQAQ 18
:|:|||||:
Db 143 RDLNTRSLRAKQQAQAQ 160

RESULT 7
F81090
hypothetical protein NMB1369 [imported] - Neisseria meningitidis (strain MC58 serogroup
C:Species: Neisseria meningitidis
C:Date: 31-Mar-2000 #sequence_revision 31-Mar-2000 #text_change 09-Jul-2004
C:Accession: F81090
R:Tetelin, H.; Saunders, N.J.; Heideberg, J.; Jeffries, A.C.; Nelson, K.E.; Eisen, J.J
; Hickey, E.K.; Haft, D.H.; Salzberg, S.L.; White, O.; Fleischmann, R.D.; Dougherty, B.A.;
ri, H.; Qin, H.; Vamathevan, J.; Gall, J.; Scarlato, V.; Maignani, V.; Pizza, M.
Science 287, 1809-1815, 2000
A:Authors: Grandi, G.; Sun, L.; Smith, H.O.; Fraser, C.M.; Moxon, E.R.; Rappuoli, R.; V
A:Title: Complete genome sequence of Neisseria meningitidis serogroup B strain MC58.
A:Reference number: A81000; MUID:20175755; PMID:10710307
A:Accession: F81090
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-184 <TET>
A:Cross-references: UNIPROT:Q9JTY6; UNIPARC:UPI00000C46C7; GB:AE002485; GB:AE002098; NII
A:Experimental source: serogroup B, strain MC58
C:Genetics:
C:Gene: NMB1369
C:Superfamily: Neisseria meningitidis probable membrane lipoprotein NMA1581

Query Match 46.9%; Score 45; DB 2; Length 184;
Best Local Similarity 55.6%; Pred. No. 7.7;
Matches 10; Conservative 3; Mismatches 5; Indels 0; Gaps 0;

QY 1 QDSKNNLTAQQAQAQVN 18
 Db 143 RDLNTRSLRAKQAQAQVN 160

RESULT 8
 JC4298
 hyaluronan receptor - mouse
 C:Species: Mus musculus (house mouse)
 C:Date: 16-Nov-1995 #sequence_revision 08-Feb-1996 #text_change 09-Jul-2004
 C:Accession: JC4298; A42925; A41923; S21586
 R:Entwistle, J.; Zhang, S.; Yang, B.; Wong, C.; Li, Q.; Hall, C.L.; A, J.; Mowat, M.; Gd
 Gene 163, 233-238, 1995
 A:Title: Characterization of the murine gene encoding the hyaluronan receptor RHAMM.
 A:Reference number: JC4298; MUID:96011639; PMID:7590272
 A:Accession: JC4298
 A:Molecule type: mRNA
 A:Residues: 1-631 <ENT>
 A:CROSS-references: UNIPROT:Q00547; UNIPARC:UPI0000154784; EMBL:X64550
 A:Experimental source: 3T3 fibroblast
 R:Hardwick, C.
 J. Cell Biol. 118, 753, 1992
 A:Reference number: A42925; MUID:92348516; PMID:1639856
 A:Contents: erratum
 A:Accession: A42925
 A>Status: preliminary
 A:Molecule type: mRNA
 A:Residues: 155-376, 'S', 378-504, 'E', 506-631 <HAR>
 A:CROSS-references: UNIPARC:UPI000014846; GB:X64550
 A:Note: authors translated the codon AGT for residue 377 as Thr and CTA for residue 507
 R:Hardwick, C.; Hoare, K.; Owens, R.; Hohn, H.P.; Hook, M.; Moore, D.; Cripps, V.; Auste
 J. Cell Biol. 117, 1343-1350, 1992
 A:Title: Molecular cloning of a novel hyaluronan receptor that mediates tumor cell motil
 A:Reference number: A41923; MUID:92299690; PMID:1376732
 A:Accession: A41923
 A>Status: preliminary
 A:Molecule type: mRNA
 A:Residues: 155-376, 'S', 378-504, 'E', 506-507, 'L', 508-630 <HAR2>
 A:CROSS-references: UNIPARC:UPI000017804; GB:X64550
 A:Note: this sequence has been corrected in reference A42925
 C:Comment: This protein regulates cell motility and transformation, and focal adhesion d
 C:Genetics:
 A:Gene: rhann
 A:Introns: 20/3; 54/2; 79/2; 137/3; 187/3; 259/2; 382/2; 431/2; 482/2; 515/3; 574/3; 625
 C:Superfamily: hyaluronan receptor
 C:Keywords: glycoprotein; receptor
 F:260-382/Region: 21 residue repeats
 F:516-574/Region: hyaluronan binding #status predicted
 F:575-625/Region: hyaluronan binding #status predicted
 F:116,283,304,325,346,367,398,438,619/Binding site: carbohydrate (Asn) (covalent) #statu

Query Match 46.9%; Score 45; DB 2; Length 631;
 Best Local Similarity 44.4%; Pred. No. 28;
 Matches 8; Conservative 5; Mismatches 5; Indels 0; Gaps 0;

QY 1 QDSKNNLTAQQAQAQVN 18
 Db 392 EDLKLNTLQSKVAAE 409

RESULT 9
 D97056
 competence COMEA protein [imported] - Clostridium acetobutylicum
 C:Species: Clostridium acetobutylicum
 C:Date: 14-Sep-2001 #sequence_revision 14-Sep-2001 #text_change 09-Jul-2004
 C:Accession: D97056
 R:Nolling, J.; Breton, G.; Onelchenko, M.V.; Markarova, K.S.; Zeng, Q.; Gibson, R.; Lee,
 J.; Dally, M.J.; Bennett, G.N.; Koonin, E.V.; Smith, D.R.
 J. Bacteriol. 193, 4823-4838, 2001
 A:Title: Genome Sequence and Comparative Analysis of the Solvent-Producing Bacterium Clo
 A:Reference number: A96900; MUID:21359325; PMID:21359325
 A:Accession: D97056

A>Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-209 <KUR>
 A:CROSS-references: UNIPROT:Q97JK6; UNIPARC:UPI00000CA14F; GB:AE001437; PIDN:AAK79239.1,
 A:Experimental source: Clostridium acetobutylicum ATCC824
 C:Genetics:
 A:Gene: CAC1268

Query Match 45.8%; Score 44; DB 2; Length 209;
 Best Local Similarity 40.0%; Pred. No. 13;
 Matches 8; Conservative 5; Mismatches 7; Indels 0; Gaps 0;

QY 1 QDSKNNLTAQQAQAQVN 20
 Db 121 KDSKVTNTSSQENTQPOQN 140

RESULT 10
 AB0827
 probable membrane protein STY2810 [imported] - Salmonella enterica subsp. enterica sero
 C:Species: Salmonella enterica subsp. enterica serovar Typhi
 A:Note: this species has also been called Salmonella typhi
 C:Date: 09-Nov-2001 #sequence_revision 09-Nov-2001 #text_change 18-Nov-2002
 C:Accession: AB0827
 R:Parkhill, J.; Dougan, G.; James, K.D.; Thomson, N.R.; Pickard, D.; Wain, J.; Churcher,
 th, T.; Conington, P.; Cronin, A.; Davis, P.; Davies, R.M.; Dowd, L.; White, N.; Farrar,
 S.; Moule, S.; O'Gaora, P.
 Nature 413, 848-852, 2001
 A:Authors: Party, C.; Quail, M.; Rutherford, K.; Simmonds, M.; Skelton, J.; Stevens, K.,
 A:Title: Complete genome sequence of a multiple drug resistant Salmonella enterica sero
 A:Reference number: AB0502; MUID:21534947; PMID:11677608
 A:Accession: AB0827
 A>Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-254 <PAR>
 A:CROSS-references: UNIPARC:UPI00000CDB90; GB:AL513382; PIDN:CAD02766.1; PID:G16503776;
 C:Genetics:
 A:Gene: STY2810

Query Match 45.8%; Score 44; DB 2; Length 254;
 Best Local Similarity 47.4%; Pred. No. 16;
 Matches 9; Conservative 5; Mismatches 5; Indels 0; Gaps 0;

QY 2 DSKNNLTAQQAQAQVN 20
 Db 182 DSELDALRQQAQALQTQLD 200

RESULT 11
 S16322
 light-induced protein CPRF-1 - parsley
 C:Species: Petroselinum crispum (parsley)
 C:Date: 13-Jan-1995 #sequence_revision 13-Jan-1995 #text_change 31-Dec-2004
 C:Accession: S16322
 R:Wesselaar, B.; Armstrong, G.A.; Block, A.; da Costa e Silva, O.; Hahlbrock, K.
 EMBO J. 10, 1777-1786, 1991
 A:Title: Light-inducible and constitutively expressed DNA-binding proteins recognizing .
 A:Reference number: S16320; MUID:91266906; PMID:2050115
 A:Accession: S16322
 A>Status: preliminary
 A:Molecule type: mRNA
 A:Residues: 1-414 <WEI>
 A:CROSS-references: UNIPROT:Q99089; UNIPARC:UPI00001794BE; EMBL:X58575
 C:Superfamily: BZIP G-box-binding protein; fos/jun DNA-binding domain homology
 C:Keywords: DNA binding; nucleus; transcription regulation
 F:267-307/Domain: fos/jun DNA-binding domain homology <FJD>

Query Match 45.8%; Score 44; DB 2; Length 414;
 Best Local Similarity 41.2%; Pred. No. 26;
 Matches 7; Conservative 7; Mismatches 3; Indels 0; Gaps 0;

QY 4 KNNLTAQQAQAQVN 20
 Db 121 KDSKVTNTSSQENTQPOQN 140

Db 303 KVDSLTAENMALKABIN 319

RESULT 12

T50010

hypothetical protein T31P16.50 - Arabidopsis thaliana

C:Species: Arabidopsis thaliana (mouse-ear cress)

C>Date: 02-Jun-2000 #sequence_revision 02-Jun-2000 #text_change 09-Jul-2004

C:Accession: T50010

R:Bevan, M.; Zimmermann, W.; Gruenewald, A.; Wambutt, R.; Kalicki, J.; Wohlmann, P.; Smith, T.; Connerton, P.; Cronin, A.; Davis, P.; Davies, R.M.; Dowd, L.; White, N.; Farrar, S.; Moule, S.; O'Gaora, P.

A:Reference number: Z55027

A:Accession: T50010

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-414 <BEV>

A:Cross-references: UNIPROT:Q9LX22; UNIPARC:UPI00000A53F5; EMBL:AL356332; GSPDB:GN000637

A:Experimental source: cultivar Columbia; BAC clone T31P16

C:Genetics:

A:Gene: ATSP:T31P16.50

A:Map position: 5

A:Introns: 55/3; 107/3; 161/1; 207/3; 219/3

Query Match 45.8%; Score 44; DB 2; Length 414;

Best Local Similarity 50.0%; Pred. No. 26;

Matches 9; Conservative 5; Mismatches 4; Indels 0; Gaps 0;

QY 1 QDSKNNLTAAQQAQVN 20

Db 207 QSELDNLKAIQVAKQ 224

RESULT 13

B86899

hypothetical protein dnah [imported] - Lactococcus lactis subsp. lactis (strain IL1403)

C:Species: Lactococcus lactis subsp. lactis

C>Date: 23-Mar-2001 #sequence_revision 23-Mar-2001 #text_change 09-Jul-2004

C:Accession: B86899

R:Botolin, A.; Wincker, P.; Mauger, S.; Jaillon, O.; Malarne, K.; Weissbach, J.; Ehrlich, G.

A:Title: The complete genome sequence of the lactic acid bacterium Lactococcus lactis ssp. cremoris

A:Reference number: A86625; MUID:21235186; PMID:11337471

A:Accession: B86899

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-552 <STO>

A:Cross-references: UNIPROT:Q9CDM5; UNIPARC:UPI00000C6BE9; GB:AE005176; PID:g12725258; F

A:Experimental source: strain IL1403

C:Genetics:

A:Gene: dnah

C:Superfamily: DNA-directed DNA polymerase III gamma chain

Query Match 45.8%; Score 44; DB 2; Length 552;

Best Local Similarity 55.6%; Pred. No. 35;

Matches 10; Conservative 2; Mismatches 6; Indels 0; Gaps 0;

QY 3 SKNNLTAAQQAQVN 20

Db 365 TEINQLKAIQVAKQ 382

RESULT 14

AD0976

probable lipoprotein STY4106 [imported] - Salmonella enterica subsp. enterica serovar Typhimurium

C:Species: Salmonella enterica subsp. enterica serovar Typhi

A>Note: this species has also been called Salmonella typhi

C>Date: 09-Nov-2001 #sequence_revision 09-Nov-2001 #text_change 18-Nov-2002

C:Accession: AD0976

R:Parkhill, J.; Dougan, G.; James, K.D.; Thomson, N.R.; Pickard, D.; Wain, J.; Churcher, C.; Connor, T.; Cronin, A.; Davis, P.; Davies, R.M.; Dowd, L.; White, N.; Farrar, S.; Moule, S.; O'Gaora, P.

A:Reference number: Z55027

A:Accession: AD0976

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-227 <PAR>

A:Cross-references: UNIPARC:UPI000005A715; GB:AL513382; PIDN:CAD03304.1; PID:g16504924;

C:Genetics:

A:Gene: STY4106

Query Match 44.8%; Score 43; DB 2; Length 227;

Best Local Similarity 40.0%; Pred. No. 20;

Matches 8; Conservative 6; Mismatches 6; Indels 0; Gaps 0;

QY 1 QDSKNNLTAAQQAQVN 20

Db 41 QEAKINLLEKQKQEVKIN 60

RESULT 15

JC1419

Fc gamma (IgG) receptor II precursor - Streptococcus sp.

N:Alternate names: fcrv protein

C:Species: Streptococcus sp.

C>Date: 30-Sep-1993 #sequence_revision 30-Sep-1993 #text_change 09-Jul-2004

C:Accession: JC1419; S17354

R:Smirnov, O.Y.; Denevsky, A.I.; Zakharov, M.V.; Abramov, V.M.; Zav'yalov, V.P.

Gene 120, 27-32, 1992

A:Title: Protein V, a novel type-II IgG receptor from Streptococcus sp.: Sequence, homology and function

A:Reference number: JC1419; MUID:93013016; PMID:1398120

A:Accession: JC1419

A:Molecule type: DNA

A:Residues: 1-587 <SMI>

A:Cross-references: UNIPROT:Q55312; UNIPARC:UPI00000BCFP7; EMBL:X62467; NID:g47562; PIDN

A:Experimental source: strain 22/58/Valente'

C:Genetics:

A:Gene: fcrv

C:Superfamily: M5 protein

C:Keywords: duplication; immunoglobulin receptor

F:1-41/Domain: signal sequence #status predicted <SIG>

F:42-587/Product: IGG Fc receptor II #status predicted <MAT>

F:234-268/Region: 35-residue repeat A

F:269-303/Region: 35-residue repeat A

F:304-339/Region: 35-residue repeat A

F:339-373/Region: 35-residue repeat A

F:374-408/Region: 35-residue repeat B

F:416-450/Region: 35-residue repeat B

Query Match 44.8%; Score 43; DB 2; Length 587;

Best Local Similarity 53.3%; Pred. No. 55;

Matches 8; Conservative 4; Mismatches 3; Indels 0; Gaps 0;

QY 4 KNNLTAAQQAQVN 18

Db 137 KVNLSAQNKALQAE 151

Search completed: January 25, 2006, 19:15:18

Job time : 10.1 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2006 Compugen Ltd.

OM protein - protein search, using sw model

Run on: January 25, 2006, 18:48:24 ; Search time 70.3 Seconds
(without alignments)
200.719 Million cell updates/sec

Title: US-10-797-821-7
Perfect score: 96
Sequence: 1 QDSKINNLTAAQQAQVN 20

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 2166443 seqs, 705528306 residues

Total number of hits satisfying chosen parameters: 2166443

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : UniProt 05.80.*

1: uniprot_sprot.*

2: uniprot_trembl.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	96	100.0	431	2 Q938V0_STRMU	Q938V0 streptococc
2	96	100.0	431	2 Q938V3_STRMU	Q938V3 streptococc
3	96	100.0	431	2 Q9AG98_STRMU	Q9AG98 streptococc
4	96	100.0	431	2 Q8DWM3_STRMU	Q8DWM3 streptococc
5	96	100.0	432	2 Q938V1_STRMU	Q938V1 streptococc
6	96	100.0	432	2 Q938V2_STRMU	Q938V2 streptococc
7	73	76.0	392	2 Q8DMY4_STRR6	Q8DMY4 streptococc
8	73	76.0	392	2 Q97N55_STRPN	Q97N55 streptococc
9	64.5	67.2	211	2 Q9ZAS7_STRMU	Q9ZAS7 streptococc
10	61	63.5	398	2 Q5XEL1_STRP6	Q5XEL1 streptococc
11	61	63.5	398	2 Q9A1Z8_STRPY	Q9A1Z8 streptococc
12	61	63.5	398	2 Q7CNQ7_STRP8	Q7CNQ7 streptococc
13	61	63.5	398	2 Q8P318_STRP3	Q8P318 streptococc
14	55	57.3	474	2 Q5M6K4_STRT2	Q5M6K4 streptococc
15	55	57.3	485	2 Q5M212_STRT1	Q5M212 streptococc
16	54	56.2	278	2 Q810C2_DROME	Q810C2 drosophila
17	54	56.2	325	2 Q9VN98_DROME	Q9VN98 drosophila
18	54	56.2	478	2 Q8SX21_DROME	Q8SX21 drosophila
19	53	55.2	4540	2 Q5SEG2_DICD1	Q5SEG2 dictyosteli
20	52	54.2	447	2 Q9AKA4_STRAG	Q9AKA4 streptococc
21	52	54.2	447	2 Q8E2H1_STRAS5	Q8E2H1 streptococc
22	52	54.2	447	2 Q8E7X9_STRAS3	Q8E7X9 streptococc
23	52	54.2	482	2 Q931K4_ENTPA	Q931K4 enterococcu
24	52	54.2	1050	2 Q70P98_9DELT	Q70P98 melittangiu
25	50.5	52.6	1189	2 Q415T6_STAHJ	Q415T6 staphylococ
26	50	52.1	343	2 Q8D5M8_VIBVU	Q8D5M8 vibrio vuln
27	50	52.1	343	2 Q7MCN2_VIBVY	Q7MCN2 vibrio vuln
28	48	50.0	123	1 KCNE2_CAVPO	P63160 cavia porce
29	48	50.0	123	1 KCNE2_RAT	P63161 rattus norv
30	48	50.0	123	2 Q5FVT9_RAT	Q5FVT9 rattus norv
31	48	50.0	516	1 P54_ENTFC	P13692 enterococcu

32	48	50.0	576	2 Q9KJJ3_ENTHR	Q9KJJ3 enterococcu
33	48	50.0	618	2 Q9VZB5_DROME	Q9VZB5 drosophila
34	48	50.0	638	1 D9NAK_PRELO	Q939f1 prevotella
35	48	50.0	1051	2 Q5MD36_9DELT	Q5MD36 cystobacter
36	47	49.0	142	2 Q83Z41_POLCB	Q83Z41 polyangium
37	47	49.0	227	2 Q57IE2_SALCH	Q57IE2 salmonella
38	47	49.0	227	2 Q8ZL65_SALTY	Q8ZL65 salmonella
39	47	49.0	391	2 Q87JM5_VIBPA	Q87JM5 vibrio para
40	47	49.0	2238	2 Q846X5_STRCM	Q846X5 streptomyce
41	47	49.0	4003	2 Q9RFL0_STIAU	Q9RFL0 stigmatella
42	46	47.9	389	1 SUCC_HAEIN	P45101 haemophilus
43	46	47.9	524	2 Q9K2J9_ENTFC	Q9K2J9 enterococcu
44	46	47.9	647	1 VE1_HPV52	P36730 human papil
45	46	47.9	940	2 Q629B0_CAEBR	Q629B0 caenorhabdi

ALIGNMENTS

```
RESULT 1
Q938V0_STRMU PRELIMINARY; PRT; 431 AA.
AC Q938V0;
DT 01-DEC-2001 (Tremblrel. 19, Created)
DT 01-DEC-2001 (Tremblrel. 19, Last sequence update)
DT 01-MAR-2004 (Tremblrel. 26, Last annotation update)
DE Glucan-binding protein B.
OS Streptococcus mutans.
OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
OX Streptococcus
NCBI_TaxID=1309;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=SSM3;
RA Jin S., Duncan M.J., Taubman M.A., Smith D.J.;
RT "Cloning of the gbpB gene from Streptococcus mutans.";
RL J. Dent. Res. 79:224-224 (2000).
RN [2]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=SSM3;
RX MEDLINE=21481971; PubMed=11598068;
EX DOI=10.1128/IAI.69.11.6931-6941.2001;
RA Maccos-Graner R.O., Jin S., King W.F., Chen T., Smith D.J.,
Duncan M.J.;
RT "Cloning of the Streptococcus mutans gene encoding glucan binding
protein B and analysis of genetic diversity and protein production in
RT clinical isolates.";
RL Infect. Immun. 69:6931-6941 (2001).
DR EMBL; AF046414; AAK94504.1; -; Genomic_DNA.
DR InterPro; IPR007921; CHAP.
DR Pfam; PF05257; CHAP; 1.
DR PRINTS; PR01852; SIBAPROTEIN.
DR PROSITE; PS50911; CHAP; 1.
SQ SEQUENCE 431 AA; 44650 MW; 05D38D8DBEC4609F CRC64;
```

Query Match 100.0%; Score 96; DB 2; Length 431;
Best Local Similarity 100.0%; Pred. No. 2.6e-06;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 QDSKINNLTAAQQAQVN 20
|||||
DB 37 QDSKINNLTAAQQAQVN 56

```
RESULT 2
Q938V3_STRMU PRELIMINARY; PRT; 431 AA.
ID Q938V3;
AC Q938V3;
DT 01-DEC-2001 (Tremblrel. 19, Created)
DT 01-DEC-2001 (Tremblrel. 19, Last sequence update)
DT 01-MAR-2004 (Tremblrel. 26, Last annotation update)
DE Glucan-binding protein B.
```

OS Streptococcus mutans.
 OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
 OC Streptococcus.
 OX NCBI_TaxID=1309;
 [1]
 RN NUCLEOTIDE SEQUENCE.
 RC STRAIN=SGJ32;
 RA Jin S., Duncan M.J., Taubman M.A., Smith D.J.;
 RT "Cloning of the gbpB gene from Streptococcus mutans.";
 RL J. Dent. Res. 79:224-224(2000).
 [2]
 RN NUCLEOTIDE SEQUENCE.
 RC STRAIN=SGJ32;
 RX MEDLINE=21481971; PubMed=11598068;
 RA DOI=10.1128/IAI.69.11.6931-6941.2001;
 RA Mattos-Graner R.O., Jin S., King W.F., Chen T., Smith D.J.,
 RA Duncan M.J.;
 RT "Cloning of the Streptococcus mutans gene encoding glucan binding
 RT protein B and analysis of genetic diversity and protein production in
 RT clinical isolates.";
 RL Infect. Immun. 69:6931-6941(2001).
 DR EMBL; AY046410; AAK94500.1; -; Genomic_DNA.
 DR InterPro; IPR007921; CHAP.
 DR InterPro; IPR009148; Siba.
 DR Pfam; PF05257; CHAP; 1.
 DR PRINTS; PR01852; SIBAPROTEIN.
 DR PROSITE; PS050911; CHAP; 1.
 SQ SEQUENCE 431 AA; 44620 MW; 464FE3B563FB7E51 CRC64;

Query Match 100.0%; Score 96; DB 2; Length 431;
 Best Local Similarity 100.0%; Pred. No. 2.6e-06; Mismatches 0; Indels 0; Gaps 0;
 Matches 20; Conservative 0;

QY 1 QDSKINNLTAQQAQAQVN 20
 |||||
 DB 37 QDSKINNLTAQQAQAQVN 56

RESULT 3
 Q9AG98 STRMU
 ID Q9AG98 STRMU PRELIMINARY; PRT; 431 AA.
 AC Q9AG98
 DT 01-JUN-2001 (TrEMBLrel. 17, Created)
 DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
 DT 01-FEB-2005 (TrEMBLrel. 29, Last annotation update)
 DE Immunodominant glycoprotein IDG-60 (Glucan-binding protein B).
 GN Name=sagA;
 OS Streptococcus mutans.
 OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
 OC Streptococcus.
 OX NCBI_TaxID=1309;
 [1]
 RN NUCLEOTIDE SEQUENCE.
 RC STRAIN=GS-5;
 RX MEDLINE=21153617; PubMed=11254612;
 RA DOI=10.1128/IAI.69.4.2493-2501.2001;
 RA Chia J.S., Lee Y.Y., Huang P.T., Chen J.Y.;
 RT "Identification of stress-responsive genes in Streptococcus mutans by
 RT differential display reverse transcription-PCR.";
 RL Infect. Immun. 69:2493-2501(2001).
 [2]
 RN NUCLEOTIDE SEQUENCE.
 RC STRAIN=GS-5;
 RX MEDLINE=21481971; PubMed=11598074;
 RA DOI=10.1128/IAI.69.11.6987-6998.2001;
 RA Chia J.S., Chang L.Y., Shun C.T., Chang Y.Y., Chen J.Y.;
 RT "A 60-kilodalton immunodominant glycoprotein is essential for cell
 RT wall integrity and the maintenance of cell shape in Streptococcus
 RT mutans.";
 RL Infect. Immun. 69:6987-6998(2001).
 [3]
 RN NUCLEOTIDE SEQUENCE.
 RC STRAIN=GS-5;

RA Chia J.-S., Chang L.-Y., Lee Y.-Y., Chen J.-Y.;
 RL Submitted (NOV-2001) to the EMBL/GenBank/DBJ databases.
 [4]
 RN NUCLEOTIDE SEQUENCE.
 RC STRAIN=3VF4;
 RA Jin S., Duncan M.J., Taubman M.A., Smith D.J.;
 RT "Cloning of the gbpB gene from Streptococcus mutans.";
 RL J. Dent. Res. 79:224-224(2000).
 [5]
 RN NUCLEOTIDE SEQUENCE.
 RC STRAIN=3VF4;
 RX MEDLINE=21481971; PubMed=11598068;
 RX DOI=10.1128/IAI.69.11.6931-6941.2001;
 RA Mattos-Graner R.O., Jin S., King W.F., Chen T., Smith D.J.,
 RA Duncan M.J.;
 RT "Cloning of the Streptococcus mutans gene encoding glucan binding
 RT protein B and analysis of genetic diversity and protein production in
 RT clinical isolates.";
 RL Infect. Immun. 69:6931-6941(2001).
 DR EMBL; AF338445; AAK08104.1; -; Genomic_DNA.
 DR EMBL; AY046411; AAK94501.1; -; Genomic_DNA.
 DR InterPro; IPR007921; CHAP.
 DR InterPro; IPR009148; Siba.
 DR Pfam; PF05257; CHAP; 1.
 DR PRINTS; PR01852; SIBAPROTEIN.
 DR PROSITE; PS050911; CHAP; 1.
 SQ SEQUENCE 431 AA; 44592 MW; 3EBE21FC5E47232E CRC64;

Query Match 100.0%; Score 96; DB 2; Length 431;
 Best Local Similarity 100.0%; Pred. No. 2.6e-06; Mismatches 0; Indels 0; Gaps 0;
 Matches 20; Conservative 0;

QY 1 QDSKINNLTAQQAQAQVN 20
 |||||
 DB 37 QDSKINNLTAQQAQAQVN 56

RESULT 4
 Q8DWM3 STRMU
 ID Q8DWM3 STRMU PRELIMINARY; PRT; 431 AA.
 AC Q8DWM3
 DT 01-MAR-2003 (TrEMBLrel. 23, Created)
 DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
 DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
 DE Putative secreted antigen GbpB/SagA; putative peptidoglycan
 DE hydrolase.
 GN Name=gbpB; OrderedLocusNames=SMU.22;
 OS Streptococcus mutans.
 OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
 OC Streptococcus.
 OX NCBI_TaxID=1309;
 [1]
 RN NUCLEOTIDE SEQUENCE.
 RC STRAIN=UA159 / ATCC 700610 / Serotype c;
 RX MEDLINE=22295063; PubMed=12397186; DOI=10.1073/pnas.172501299;
 RA Ajdic D.J., McShan W.M., McLaughlin R.E., Savic G., Chang J.,
 RA Carlson M.B., Briceaux C., Tian R., Kenton S., Jia H.G., Lin S.P.,
 RA Qian Y., Li S., Zhu H., Najjar F.Z., Lai H., White J., Roe B.A.,
 RA Ferretti J.J.;
 RT "Genome sequence of Streptococcus mutans UA159, a cariogenic dental
 RT pathogen.";
 RL Proc. Natl. Acad. Sci. U.S.A. 99:14434-14439(2002).
 DR EMBL; AE014855; AAN57811.1; -; Genomic_DNA.
 DR InterPro; IPR007921; CHAP.
 DR InterPro; IPR009148; Siba.
 DR Pfam; PF05257; CHAP; 1.
 DR PRINTS; PR01852; SIBAPROTEIN.
 DR PROSITE; PS050911; CHAP; 1.
 KW Complete proteome.
 SQ SEQUENCE 431 AA; 44620 MW; 2D1CA685248CCD3E CRC64;

Query Match 100.0%; Score 96; DB 2; Length 431;
 Best Local Similarity 100.0%; Pred. No. 2.6e-06; Mismatches 0; Indels 0; Gaps 0;

Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 QDSKINLTAQQAQAQVN 20
 |||||
 DB 37 QDSKINLTAQQAQAQVN 56

RESULT 5

Q938V1_STRMU PRELIMINARY; PRT; 432 AA.
 AC Q938V1;
 DT 01-DEC-2001 (TrEMBLrel. 19, Created)
 DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
 DT 01-WAR-2004 (TrEMBLrel. 26, Last annotation update)
 DE Glucan-binding protein B.
 OS Streptococcus mutans.
 OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
 OC Streptococcus.
 OX NCBI_TaxID=1309;
 RN [1]_TaxID=1309;
 RP NUCLEOTIDE SEQUENCE.
 RC STRAIN=3SN1;
 RA Jin S., Duncan M.J., Taubman M.A., Smith D.J.;
 RT "Cloning of the gbpB gene from Streptococcus mutans.";
 RL J. Dent. Res. 79:224-224(2000).
 RN [2]

NUCLEOTIDE SEQUENCE.

RC STRAIN=3SN1;
 RX MEDLINE=21481971; PubMed=11598068;
 RX DOI=10.1128/IAI.69.11.6931-6941.2001;
 RA Mattos-Graner R.O., Jin S., King W.F., Chen T., Smith D.J.,
 Duncan M.J.;
 RA "Cloning of the Streptococcus mutans gene encoding glucan binding
 RT protein B and analysis of genetic diversity and protein production in
 RT clinical isolates.";
 RL Infect. Immun. 69:6931-6941(2001).
 DR EMBL: AY046413; AAK94503.1; -; Genomic_DNA.
 DR InterPro; IPR007921; CHAP.
 DR InterPro; IPR009148; Siba.
 DR Pfam; PF05257; CHAP; 1.
 DR PRINTS; PR01852; SIBAPROTEIN.
 DR PROSITE; PS50911; CHAP; 1.
 SQ SEQUENCE 432 AA; 44652 MW; 3F88ECB9A1F3BE4F CRC64;

Query Match 100.0%; Score 96; DB 2; Length 432;
 Best Local Similarity 100.0%; Pred. No. 2.6e-06;
 Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 QDSKINLTAQQAQAQVN 20
 |||||
 DB 37 QDSKINLTAQQAQAQVN 56

RESULT 6

Q938V2_STRMU PRELIMINARY; PRT; 432 AA.
 AC Q938V2;
 DT 01-DEC-2001 (TrEMBLrel. 19, Created)
 DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
 DT 01-WAR-2004 (TrEMBLrel. 26, Last annotation update)
 DE Glucan-binding protein B.
 OS Streptococcus mutans.
 OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
 OC Streptococcus.
 OX NCBI_TaxID=1309;
 RN [1]_TaxID=1309;
 RP NUCLEOTIDE SEQUENCE.
 RC STRAIN=15JP2;
 RA Jin S., Duncan M.J., Taubman M.A., Smith D.J.;
 RT "Cloning of the gbpB gene from Streptococcus mutans.";
 RL J. Dent. Res. 79:224-224(2000).
 RN [2]

NUCLEOTIDE SEQUENCE.

RC STRAIN=15JP2;
 RX MEDLINE=21481971; PubMed=11598068;
 RX DOI=10.1128/IAI.69.11.6931-6941.2001;
 RA Mattos-Graner R.O., Jin S., King W.F., Chen T., Smith D.J.,
 Duncan M.J.;
 RA "Cloning of the Streptococcus mutans gene encoding glucan binding
 RT protein B and analysis of genetic diversity and protein production in
 RT clinical isolates.";
 RL Infect. Immun. 69:6931-6941(2001).
 DR EMBL: AY046412; AAK94502.1; -; Genomic_DNA.
 DR InterPro; IPR007921; CHAP.
 DR InterPro; IPR009148; Siba.
 DR Pfam; PF05257; CHAP; 1.
 DR PRINTS; PR01852; SIBAPROTEIN.
 DR PROSITE; PS50911; CHAP; 1.
 SQ SEQUENCE 432 AA; 44648 MW; E769B2504AE50E9 CRC64;

Query Match 100.0%; Score 96; DB 2; Length 432;
 Best Local Similarity 100.0%; Pred. No. 2.6e-06;
 Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 QDSKINLTAQQAQAQVN 20
 |||||
 DB 37 QDSKINLTAQQAQAQVN 56

RESULT 7

Q8DMY4_STRR6 PRELIMINARY; PRT; 392 AA.
 AC Q8DMY4;
 DT 01-WAR-2003 (TrEMBLrel. 23, Created)
 DT 01-WAR-2003 (TrEMBLrel. 23, Last sequence update)
 DT 01-WAR-2004 (TrEMBLrel. 26, Last annotation update)
 DE General stress protein GSP-781.
 GN Name=gsp-781; OrderedLocNames=spr2021;
 OS Streptococcus pneumoniae (strain ATCC BAA-255 / R6).
 OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
 OC Streptococcus.
 OX NCBI_TaxID=171101;
 RN [1]
 RP NUCLEOTIDE SEQUENCE.
 RX MEDLINE=21429245; PubMed=11544234;
 RX DOI=10.1128/JB.183.19.5709-5717.2001;
 RA Hoskins J., Alborn W.E. Jr., Arnold J., Blaszcak L.C., Burgett S.,
 DeHoff B.S., Estrem S.T., Fritz L., Fu D.-J., Fuller W., Geringer C.,
 Gilmore R., Glass J.S., Khoja H., Kraft A.R., Lagace R.E.,
 LeBlanc D.J., Lee L.N., Lefkowitz E.J., Lu J., Matsushima P.,
 McAhren S.M., McHenry M., McLeaster K., Mundy C.W., Nicas T.I.,
 Norris F.H., O'Gara M., Peery R.B., Robertson G.T., Rockey P.,
 Sun P.-M., Winkler M.E., Yang Y., Young-Bellido M., Zhao G.,
 Zook C.A., Baltz R.H., Jaskunas S.R., Rostek P.R. Jr., Skatrud P.L.,
 Glass J.I.;
 RA "Genome of the bacterium Streptococcus pneumoniae strain R6.";
 RL J. Bacteriol. 183:5709-5717(2001).
 DR EMBL: AE008566; AAL00823.1; -; Genomic_DNA.
 DR PIR; B98124; B98124.
 DR PIR; G95258; G95258.
 DR InterPro; IPR007921; CHAP.
 DR InterPro; IPR009148; Siba.
 DR Pfam; PF05257; CHAP; 1.
 DR PRINTS; PR01852; SIBAPROTEIN.
 DR PROSITE; PS50911; CHAP; 1.
 KW Complete proteome.
 SQ SEQUENCE 392 AA; 41697 MW; B5A5A860FAEA77DD CRC64;

Query Match 76.0%; Score 73; DB 2; Length 392;
 Best Local Similarity 75.0%; Pred. No. 0.0085;
 Matches 15; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

QY 1 QDSKINLTAQQAQAQVN 20
 |||||
 DB 37 QDNKSLNLTAAQQAQAQVN 56

```

Matches 16; Conservative 1; Mismatches 3; Indels 7; Gaps 1;

QY 1 QDSKINNLTAAQQ-----AAQAQVN 20
Db 37 QDSKINNLTAAQQQHHKHLITIQQVS 63

RESULT 10
Q9XEL1_STRP6
ID Q9XEL1_STRP6 PRELIMINARY; PRT; 398 AA.
AC Q9XEL1;
DT 25-OCT-2004 (TrEMBLrel. 28, Created)
DT 25-OCT-2004 (TrEMBLrel. 28, Last sequence update)
DE 25-OCT-2004 (TrEMBLrel. 28, Last annotation update)
DE Secreted protein.
GN OrderedLocusNames=M6_Spy0017;
OS Streptococcus pyogenes (serotype M6).
OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
OC Streptococcus.
OX NCBI_TaxID=301450;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=MGAS10394;
RX PubMed=15272401; DOI=10.1086/422697;
RA Banks D.J., Porcella S.F., Barbican K.D., Beres S.B., Phillips L.B.,
RA Voyich J.M., DeLeo P.R., Martin J.M., Somerville G.A., Musser J.M.;
RT "Progress toward characterization of the group A Streptococcus
metagenome: complete genome sequence of a macrolide-resistant serotype
M6 strain.";
RL J. Infect. Dis. 190:727-738(2004).
DR EMBL; CP000003; AAT86152.1; -; Genomic_DNA.
DR InterPro; IPR007921; CHAP.
DR InterPro; IPR009148; Siba.
DR Pfam; PF05257; CHAP; 1.
DR PRINTS; PR01852; SIBAPROTEIN.
DR PROSITE; PS0911; CHAP; 1.
KW Complete proteome.
SQ SEQUENCE 398 AA; 42028 MW; 5DCEDA78CB863B60 CRC64;

Query Match 63.5%; Score 61; DB 2; Length 398;
Best Local Similarity 65.0%; Pred. No. 0.62;
Matches 13; Conservative 4; Mismatches 3; Indels 0; Gaps 0;

QY 1 QDSKINNLTAAQQAAQAVN 20
Db 34 QDSIIISLTTTQKAAQNV 53

RESULT 11
Q9A1Z8_STRPY
ID Q9A1Z8_STRPY PRELIMINARY; PRT; 398 AA.
AC Q9A1Z8; Q7BH59;
DT 01-JUN-2001 (TrEMBLrel. 17, Created)
DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
DT 10-MAY-2005 (TrEMBLrel. 30, Last annotation update)
DE Secreted protein Siba precursor.
GN OrderedLocusNames=SPY0019;
OS Streptococcus pyogenes.
OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
OC Streptococcus.
OX NCBI_TaxID=1314;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=SF370 / ATCC 700294 / Serotype M1;
RX MEDLINE=21192684; PubMed=11296296; DOI=10.1073/pnas.071559398;
RA Ferretti J.J., McShan W.M., Ajdic D.J., Savic D.J., Savic G., Lyon K.,
RA Primeaux C., Sezate S., Suvorov A.N., Kenton S., Lai H.S., Lin S.P.,
RA Qian Y., Jia H.G., Najjar F.Z., Ren Q., Zhu H., Song L., White J.,
RA Yuan X., Clifton S.W., Roe B.A., McLaughlin R.E.;
RT "Complete genome sequence of an M1 strain of Streptococcus pyogenes.";
RL Proc. Natl. Acad. Sci. U.S.A. 98:4658-4663(2001).
RN [2]
RP NUCLEOTIDE SEQUENCE.

Matches 15; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

QY 1 QDSKINNLTAAQQAAQAVN 20
Db 37 QDNKISLNTAAQQEAQKQVD 56

RESULT 9
Q9ZAS7_STRMU
ID Q9ZAS7_STRMU PRELIMINARY; PRT; 211 AA.
AC Q9ZAS7;
DT 01-MAY-1999 (TrEMBLrel. 10, Created)
DT 01-MAY-1999 (TrEMBLrel. 10, Last sequence update)
DT 01-OCT-2002 (TrEMBLrel. 22, Last annotation update)
DE Putative secreted protein (fragment).
OS Streptococcus mutans.
OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
OC Streptococcus.
OX NCBI_TaxID=1309;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=GS-5 Kuramitsu;
RA Peruzzi P., Piggot P.J., Daneo-Moore L.;
RL Submitted (NOV-1996) to the EMBL/GenBank/DBJ databases.
DR EMBL; U78607; AAD00288.1; -; Genomic_DNA.
FT NON_TER 211
SQ SEQUENCE 211 AA; 22803 MW; 4ACE331159CFAFC6 CRC64;

Query Match 67.2%; Score 64.5; DB 2; Length 211;
Best Local Similarity 59.3%; Pred. No. 0.096;

```

```

RA Fagan P.K., Reinscheid D., Gottschalk B., Chhatwal G.S.;
RT "Identification and characterization of a novel secreted protein from
RT group A streptococcus.";
RL Submitted (NOV-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AR006474; AAK3158.1; -; Genomic DNA.
DR EMBL; AF319999; AAL73135.1; -; Genomic DNA.
DR InterPro; IPR007921; CHAP.
DR InterPro; IPR009148; Siba.
DR Pfam; PF05257; CHAP; 1.
DR PRINTS; PR01852; SIBAPROTEIN.
DR PROSITE; PS0911; CHAP; 1.
KW Complete proteome; Signal.
FT SIGNAL 1 23 Potential
SQ SEQUENCE 398 AA; 41899 MW; 28A9B3F7195E969B CRC64;

Query Match 63.5%; Score 61; DB 2; Length 398;
Best Local Similarity 65.0%; Pred. No. 0.62;
Matches 13; Conservative 4; Mismatches 3; Indels 0; Gaps 0;

QY 1 QDSKINNLTAAQQAAQAQVN 20
DB 34 QDSIISNLTTQKAAQNVQS 53

RESULT 12
ID Q7CNQ7_STRP8 PRELIMINARY; PRT; 398 AA.
AC Q7CNQ7;
DT 05-JUL-2004 (TrEMBLrel. 27, Created)
DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE Putative secreted protein.
GN OrderedLocusNames=spym18_0020;
OS Streptococcus pyogenes (serotype M18).
OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
OC Streptococcus.
OX NCBI_TaxID=301451;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=MGAS8232 / Serotype M18;
RX MEDLINE=21927593; PubMed=11917108; DOI=10.1073/pnas.062526099;
RA Smoot J.C., Barbican K.D., Van Gompel J.J., Smoot L.M., Chaussee M.S.,
RA Sylva G.L., Sturdevant D.E., Ricklefs S.W., Porcella S.F.,
RA Parkins L.D., Beres S.B., Campbell D.S., Smith T.M., Zhang Q.,
RA Kapur V., Daly J.A., Veasy L.G., Musser J.M.;
RA "Genome sequence and comparative microarray analysis of serotype M18
RT group A Streptococcus strains associated with acute rheumatic fever
RT outbreaks";
RL Proc. Natl. Acad. Sci. U.S.A. 99:4668-4673 (2002).
DR EMBL; AE009955; AAL96849.1; -; Genomic DNA.
DR InterPro; IPR007921; CHAP.
DR InterPro; IPR009148; Siba.
DR Pfam; PF05257; CHAP; 1.
DR PRINTS; PR01852; SIBAPROTEIN.
DR PROSITE; PS0911; CHAP; 1.
KW Complete proteome.
SQ SEQUENCE 398 AA; 42028 MW; 5DCEDA78CB863B60 CRC64;

Query Match 63.5%; Score 61; DB 2; Length 398;
Best Local Similarity 65.0%; Pred. No. 0.62;
Matches 13; Conservative 4; Mismatches 3; Indels 0; Gaps 0;

QY 1 QDSKINNLTAAQQAAQAQVN 20
DB 34 QDSIISNLTTQKAAQNVQS 53

RESULT 13
ID Q8P318_STRP3 PRELIMINARY; PRT; 398 AA.
AC Q8P318_07CFL7;
DT 01-OCT-2002 (TrEMBLrel. 22, Created)
DT 01-OCT-2002 (TrEMBLrel. 22, Last sequence update)

```

```

DT 13-SEP-2005 (TrEMBLrel. 31, Last annotation update)
DE Putative secreted protein.
GN OrderedLocusNames=Spa0015, Spym3_0014;
OS Streptococcus pyogenes (serotype M3).
OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
OC Streptococcus.
OX NCBI_TaxID=301448;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=SSI-1 / Serotype M3;
RX MEDLINE=22683278; PubMed=12799345; DOI=10.1101/gr.1096703;
RA Nakagawa I., Kurokawa K., Yamaashita A., Nakata M., Tomiyasu Y.,
RA Okahashi N., Kawabata S., Yamazaki K., Shiba T., Yasunaga T.,
RA Hayashi H., Hattori M., Hamada S.;
RA "Genome sequence of an M3 strain of Streptococcus pyogenes reveals a
RT large-scale genomic rearrangement in invasive strains and new insights
RT into phage evolution.";
RL Genome Res. 13:1042-1055 (2003).
RN [2]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=MGAS315 / Serotype M3;
RX MEDLINE=22133808; PubMed=12122206; DOI=10.1073/pnas.152298499;
RA Beres S.B., Sylva G.L., Barbican K.D., Lei B., Hoff J.S.,
RA Mammarella N.D., Liu M.-Y., Smoot J.C., Porcella S.F., Parkins L.D.,
RA Campbell D.S., Smith T.M., McCormick J.K., Leung D.Y.M.,
RA Schlievert P.M., Musser J.M.;
RA "Genome sequence of a serotype M3 strain of group A Streptococcus:
RT phage-encoded toxins, the high-virulence phenotype, and clone
RT emergence.";
RL Proc. Natl. Acad. Sci. U.S.A. 99:10078-10083 (2002).
DR EMBL; BA000034; BAC63110.1; -; Genomic DNA.
DR EMBL; AE014136; AAM78621.1; -; Genomic DNA.
DR InterPro; IPR007921; CHAP.
DR InterPro; IPR009148; Siba.
DR Pfam; PF05257; CHAP; 1.
DR PRINTS; PR01852; SIBAPROTEIN.
DR PROSITE; PS0911; CHAP; 1.
KW Complete proteome.
SQ SEQUENCE 398 AA; 42028 MW; 5DCEDA78CB863B60 CRC64;

Query Match 63.5%; Score 61; DB 2; Length 398;
Best Local Similarity 65.0%; Pred. No. 0.62;
Matches 13; Conservative 4; Mismatches 3; Indels 0; Gaps 0;

QY 1 QDSKINNLTAAQQAAQAQVN 20
DB 34 QDSIISNLTTQKAAQNVQS 53

RESULT 14
ID Q5M6K4_STRT2 PRELIMINARY; PRT; 474 AA.
AC Q5M6K4;
DT 01-FEB-2005 (TrEMBLrel. 29, Created)
DT 01-FEB-2005 (TrEMBLrel. 29, Last sequence update)
DT 13-SEP-2005 (TrEMBLrel. 31, Last annotation update)
DE Glucan binding protein (PcseB).
GN Name=pcseB; OrderedLocusNames=stu0022;
OS Streptococcus thermophilus (strain ATCC BAA-250 / LMG 18311).
OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
OC Streptococcus.
OX NCBI_TaxID=264199;
RN [1]
RP NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
RX PubMed=15543133; DOI=10.1038/nbt1034;
RA Bolotin A., Quinquis B., Renault P., Sorokin A., Ehrlich S.D.,
RA Kulakauskas S., Lapidus A., Goldsman E., Mazur M., Pusch G.D.,
RA Fonstein M., Overbeek R., Kyprides N., Purnelle B., Prozzi D.,
RA Ngui K., Masuy D., Hancy F., Burteau S., Boutry M., Delcour J.,
RA Goffeau A., Hols P.;
RA "Complete sequence and comparative genome analysis of the dairy
RT bacterium Streptococcus thermophilus.";
RT Nat. Biotechnol. 22:1554-1558 (2004).

```

```
RN [2]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=IMG 18311;
RA Borges F., Layec S., Thibessard A., Fernandez A., Gintz B., Hols P.,
RA Decaris B., Leblond-Bourget N.;
RT "cse, a Chimeric and Variable Gene, Encodes an Extracellular Protein
RT Involved in Cellular Segregation in Streptococcus thermophilus.";
RL J. Bacteriol. 187:2737-2746(2005).
DR EMBL; CP000023; AAV59752.1; -; Genomic_DNA.
DR EMBL; AY730643; AAW82375.1; -; Genomic_DNA.
DR InterPro; IPR007921; CHAP.
DR InterPro; IPR009148; Siba.
DR Pfam; PF05257; CHAP; 1.
DR PRINTS; PR01852; SIBAPROTEIN.
DR PROSITE; PS0911; CHAP; 1.
KW Complete proteome.
SQ SEQUENCE 474 AA; 48142 MW; 1BAC6A9E0A0A200D CRC64;

Query Match 57.3%; Score 55; DB 2; Length 474;
Best Local Similarity 61.1%; Pred. No. 6.2;
Matches 11; Conservative 5; Mismatches 2; Indels 0; Gaps 0;

QY 2 DSKNNLTAAQQAAQAV 19
   ::|||::|||
DB 35 NNAISNLASQQAQAV 52

RESULT 15
QSM212 STRT1
ID Q5M212_STRT1 PRELIMINARY; PRT; 485 AA.
AC QSM212;
DT 01-FEB-2005 (TrEMBLrel. 29, Created)
DT 01-FEB-2005 (TrEMBLrel. 29, Last sequence update)
DT 01-FEB-2005 (TrEMBLrel. 29, Last annotation update)
DE Glucan binding protein.
DE
GN Name=pcgB; OrderedLocustNames=str0022;
OS Streptococcus thermophilus (strain CNRZ 1066).
OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
OC Streptococcus.
OX NCBI_TaxID=299768;
RN [1]
RP NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
RX PubMed:1543133; DOI=10.1038/nbt1034;
RA Bolotin A., Quinquis B., Renault P., Sorokin A., Ehrlich S.D.,
RA Kulakauskas S., Lapidus A., Goldsman E., Mazur M., Pusch G.D.,
RA Fonstein M., Overbeek R., Kyprides N., Purnelle B., Prozzi D.,
RA Ngui K., Masuy D., Hancy F., Burteau S., Boutry M., Delcour J.,
RA Goffeau A., Hols P.;
RT "Complete sequence and comparative genome analysis of the dairy
RT bacterium Streptococcus thermophilus.";
RL Nat. Biotechnol. 22:1554-1558(2004).
DR EMBL; CP000024; AAV61641.1; -; Genomic_DNA.
DR InterPro; IPR007921; CHAP.
DR InterPro; IPR009148; Siba.
DR Pfam; PF05257; CHAP; 1.
DR PRINTS; PR01852; SIBAPROTEIN.
DR PROSITE; PS0911; CHAP; 1.
KW Complete proteome.
SQ SEQUENCE 485 AA; 49165 MW; 8665E4B1FC6CF7CA CRC64;

Query Match 57.3%; Score 55; DB 2; Length 485;
Best Local Similarity 61.1%; Pred. No. 6.3;
Matches 11; Conservative 5; Mismatches 2; Indels 0; Gaps 0;

QY 2 DSKNNLTAAQQAAQAV 19
   ::|||::|||
DB 35 NNAISNLASQQAQAV 52
```

Search completed: January 25, 2006, 19:13:39
Job time : 70.3 secs

GenCore version 5.1.1.6
Copyright (c) 1993 - 2006 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: January 25, 2006, 18:50:05 ; Search time 16.6 Seconds
(without alignments)
99.609 Million cell updates/sec

Title: US-10-797-821-7
Perfect score: 96
Sequence: 1 QDSKINLTAQQQAAQVN 20

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 572060 seqs, 82675679 residues

Total number of hits satisfying chosen parameters: 572060

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued Patents AA:*
1: /cgn2_6/ptodata/1/iaa/5 COMB.pap.*
2: /cgn2_6/ptodata/1/iaa/6 COMB.pap.*
3: /cgn2_6/ptodata/1/iaa/H COMB.pap.*
4: /cgn2_6/ptodata/1/iaa/PCTUS COMB.pap.*
5: /cgn2_6/ptodata/1/iaa/RE COMB.pap.*
6: /cgn2_6/ptodata/1/iaa/backfiles1.pap.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	73	76.0	210	2 US-09-222-938A-67	Sequence 67, Appl
2	73	76.0	392	2 US-09-583-110-4374	Sequence 4374, Ap
3	73	76.0	399	2 US-09-107-433-3230	Sequence 3230, Ap
4	54	56.2	168	2 US-09-270-767-31753	Sequence 31753, A
5	54	56.2	168	2 US-09-270-767-46970	Sequence 46970, A
6	54	56.2	325	2 US-09-972-784-4	Sequence 4, Appli
7	54	56.2	478	2 US-09-738-945-6	Sequence 6, Appli
8	52	54.2	497	2 US-09-134-000C-5990	Sequence 5990, Ap
9	48	50.0	123	2 US-09-550-163-4	Sequence 4, Appli
10	46	47.9	525	2 US-09-107-532A-5095	Sequence 5095, Ap
11	45	46.9	183	2 US-08-477-831C-41	Sequence 41, Appl
12	45	46.9	221	2 US-09-978-309A-75	Sequence 75, Appl
13	45	46.9	333	2 US-09-978-309A-73	Sequence 73, Appl
14	45	46.9	435	2 US-09-978-309A-80	Sequence 80, Appl
15	45	46.9	476	2 US-09-978-309A-77	Sequence 77, Appl
16	45	46.9	476	1 US-09-978-309A-79	Sequence 79, Appl
17	45	46.9	477	1 US-08-402-217A-3	Sequence 3, Appli
18	45	46.9	477	1 US-08-700-178-3	Sequence 3, Appli
19	45	46.9	477	2 US-08-995-654-3	Sequence 3, Appli
20	45	46.9	486	2 US-09-540-236-3513	Sequence 3513, Ap
21	45	46.9	606	2 US-08-477-831C-2	Sequence 2, Appli
22	45	46.9	631	2 US-08-477-831C-11	Sequence 11, Appl
23	45	46.9	631	2 US-09-685-010-48	Sequence 48, Appl
24	45	46.9	631	2 US-09-978-309A-48	Sequence 48, Appl
25	44	45.8	36	2 US-09-092-315-20	Sequence 20, Appl
26	44	45.8	36	2 US-09-733-524A-20	Sequence 20, Appl
27	44	45.8	36	2 US-10-189-977A-20	Sequence 20, Appl

28	44	45.8	282	2 US-09-605-703B-894	Sequence 894, App
29	44	45.8	313	2 US-09-605-703B-892	Sequence 892, App
30	43.5	45.3	885	2 US-09-710-279-1660	Sequence 1660, Ap
31	43.5	45.3	1211	2 US-09-134-001C-4820	Sequence 4820, Ap
32	42	43.8	123	2 US-09-949-016-6592	Sequence 6592, Ap
33	42	43.8	123	2 US-09-550-163-2	Sequence 2, Appli
34	42	43.8	138	2 US-09-949-016-11054	Sequence 11054, A
35	42	43.8	334	2 US-09-134-000C-6029	Sequence 6029, Ap
36	42	43.8	622	2 US-09-134-001C-4733	Sequence 4733, Ap
37	42	43.8	606	2 US-09-538-092-689	Sequence 689, App
38	42	43.8	7831	2 US-09-902-540-12902	Sequence 12902, A
39	41.5	43.2	119	2 US-09-732-210-698	Sequence 688, App
40	41	42.7	183	2 US-08-477-831C-40	Sequence 40, Appl
41	41	42.7	293	2 US-09-605-703B-884	Sequence 884, App
42	41	42.7	293	2 US-09-605-703B-886	Sequence 886, App
43	41	42.7	305	2 US-09-248-796A-19372	Sequence 19372, A
44	41	42.7	479	2 US-09-540-236-2292	Sequence 2292, Ap
45	41	42.7	523	2 US-09-489-039A-12885	Sequence 12885, A

ALIGNMENTS

RESULT 1
US-09-222-938A-67
; Sequence 67, Application US/09222938A
; Patent No. 6437108
; GENERAL INFORMATION:
; APPLICANT: Youngman, Philip
; APPLICANT: Pritz, Christian
; APPLICANT: Murphy, Christopher
; APPLICANT: Guzman, Luz-Maria
; TITLE OF INVENTION: ESSENTIAL BACTERIAL GENES AND THEIR USE
; FILE REFERENCE: 07334/060001
; CURRENT APPLICATION NUMBER: US/09/222,938A
; CURRENT FILING DATE: 1998-12-30
; NUMBER OF SEQ ID NOS: 102
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 67
; LENGTH: 210
; TYPE: PRT
; ORGANISM: Streptococcus pneumoniae
US-09-222-938A-67

Query Match 76.0%; Score 73; DB 2; Length 210;
Best Local Similarity 75.0%; Pred. No. 0.00037;
Matches 15; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

QY 1 QDSKINLTAQQQAAQVN 20
|||:|||||
DB 37 QDNKISNLTAQQQAAQKQVD 56

RESULT 2
US-09-583-110-4374
; Sequence 4374, Application US/09583110
; Patent No. 6699703
; GENERAL INFORMATION:
; APPLICANT: Lynn Doucette-Stamm et al.
; TITLE OF INVENTION: Nucleic Acid and Amino Acid Sequences Relating to Streptococcus
; FILE REFERENCE: PATH00-07A
; CURRENT APPLICATION NUMBER: US/09/583,110
; CURRENT FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: US 09/107,433
; PRIOR FILING DATE: 1998-06-30
; PRIOR APPLICATION NUMBER: US 60/085,131
; PRIOR FILING DATE: 1998-05-12
; PRIOR APPLICATION NUMBER: US 60/051,553
; PRIOR FILING DATE: 1997-07-02
; NUMBER OF SEQ ID NOS: 5322
; SEQ ID NO 4374
; LENGTH: 392

```
; TYPE: PRT
; ORGANISM: Streptococcus pneumoniae
US-09-583-110-4374

Query Match          76.0%; Score 73; DB 2; Length 392;
Best Local Similarity 75.0%; Pred. No. 0.00077;
Matches 15; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

QY 1 QDSKINNLTAAQQAQVN 20
   ||:|||||:|||||:|
Db 37 QDNKISNLTAAQQAQKQVD 56

RESULT 3
US-09-107-433-3230
; Sequence 3230, Application US/09107433
; Patent No. 6800744
; GENERAL INFORMATION:
; APPLICANT: Lynn A Doucette-Stamm and David Bush
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID
; SEQUENCES RELATING TO STREPTOCOCCUS PNEUMONIAE
; THERAPEUTICS
; NUMBER OF SEQUENCES: 5206
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: GENOME THERAPEUTICS CORPORATION
; STREET: 100 Beaver Street
; CITY: Waltham
; STATE: Massachusetts
; COUNTRY: USA
; ZIP: 02354
; MEDIUM TYPE: CD-ROM ISO9660
; COMPUTER: <Unknown>
; OPERATING SYSTEM: <Unknown>
; SOFTWARE: <Unknown>
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/107,433
; FILING DATE: 30-Jun-1998
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 60/ 085131
; FILING DATE: May 12, 1998
; APPLICATION NUMBER: 60/051553
; FILING DATE: July 2, 1997
; ATTORNEY/AGENT INFORMATION:
; NAME: Arinello, Pamela Deneke
; REGISTRATION NUMBER: 40,489
; REFERENCE/DOCKET NUMBER: GTC-011
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (781)893-5007
; TELEFAX: (781)893-8277
; INFORMATION FOR SEQ ID NO: 3230:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 399 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; HYPOTHETICAL: YES
; ORIGINAL SOURCE:
; ORGANISM: Streptococcus pneumoniae
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (B) LOCATION 1...399
; SEQUENCE DESCRIPTION: SEQ ID NO: 3230:
US-09-107-433-3230

Query Match          76.0%; Score 73; DB 2; Length 399;
Best Local Similarity 75.0%; Pred. No. 0.00078;
Matches 15; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

QY 1 QDSKINNLTAAQQAQVN 20
   ||:|||||:|||||:|
Db 44 QDNKISNLTAAQQAQKQVD 63

FOR DIAGN
```

```
RESULT 4
US-09-270-767-31753
; Sequence 31753, Application US/09270767
; Patent No. 6703491
; GENERAL INFORMATION:
; APPLICANT: Homburger et al.
; TITLE OF INVENTION: Nucleic acids and proteins of Drosophila melanogaster
; FILE REFERENCE: File Reference: 7326-094
; CURRENT APPLICATION NUMBER: US/09/270,767
; CURRENT FILING DATE: 1999-03-17
; NUMBER OF SEQ ID NOS: 62517
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 31753
; LENGTH: 168
; TYPE: PRT
; ORGANISM: Drosophila melanogaster
US-09-270-767-31753

Query Match          56.2%; Score 54; DB 2; Length 168;
Best Local Similarity 61.1%; Pred. No. 0.33;
Matches 11; Conservative 3; Mismatches 4; Indels 0; Gaps 0;

QY 1 QDSKINNLTAAQQAQ 18
   :||| ||| ||| |||
Db 105 ENSKTNLAAQAQAQAE 122

RESULT 5
US-09-270-767-46970
; Sequence 46970, Application US/09270767
; Patent No. 6703491
; GENERAL INFORMATION:
; APPLICANT: Homburger et al.
; TITLE OF INVENTION: Nucleic acids and proteins of Drosophila melanogaster
; FILE REFERENCE: File Reference: 7326-094
; CURRENT APPLICATION NUMBER: US/09/270,767
; CURRENT FILING DATE: 1999-03-17
; NUMBER OF SEQ ID NOS: 62517
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 46970
; LENGTH: 168
; TYPE: PRT
; ORGANISM: Drosophila melanogaster
US-09-270-767-46970

Query Match          56.2%; Score 54; DB 2; Length 168;
Best Local Similarity 61.1%; Pred. No. 0.33;
Matches 11; Conservative 3; Mismatches 4; Indels 0; Gaps 0;

QY 1 QDSKINNLTAAQQAQ 18
   :||| ||| ||| |||
Db 105 ENSKTNLAAQAQAQAE 122

RESULT 6
US-09-972-784-4
; Sequence 4, Application US/09972784
; Patent No. 6566088
; GENERAL INFORMATION:
; APPLICANT: McKnight, Steven L.
; APPLICANT: Bruik, Richard K.
; TITLE OF INVENTION: Prolyl-4-Hydroxylases
; FILE REFERENCE: UTSD0871
; CURRENT APPLICATION NUMBER: US/09/972,784
; CURRENT FILING DATE: 2001-10-04
; NUMBER OF SEQ ID NOS: 5
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 4
; LENGTH: 325
; TYPE: PRT
; ORGANISM: Drosophila melanogaster
US-09-972-784-4
```


Query Match 56.2%; Score 54; DB 2; Length 325;
 Best Local Similarity 61.1%; Pred. No. 0.7;
 Matches 11; Conservative 3; Mismatches 4; Indels 0; Gaps 0;

QY 1 QDSKNNLTAQQQAQAQ 18
 DB 262 ENSKTNLAAQAQAQAE 279

RESULT 7
 US-09-738-946-6
 ; Sequence 6, Application US/09738946
 ; Patent No. 6579701
 ; GENERAL INFORMATION:
 ; APPLICANT: EXELIXIS, INC.
 ; TITLE OF INVENTION: DROSOPHILA HOMOLOGUES OF GENES AND PROTEINS IMPLICATED IN CANCER
 ; FILE REFERENCE: EX00-043C
 ; CURRENT APPLICATION NUMBER: US/09/738,946
 ; CURRENT FILING DATE: 2000-12-14
 ; PRIOR APPLICATION NUMBER: 60/170,832
 ; PRIOR FILING DATE: 1999-12-14
 ; PRIOR APPLICATION NUMBER: 60/170,838
 ; PRIOR FILING DATE: 1999-12-14
 ; PRIOR APPLICATION NUMBER: 60/178,580
 ; PRIOR FILING DATE: 2000-01-28
 ; PRIOR APPLICATION NUMBER: 60/185,879
 ; PRIOR FILING DATE: 2000-02-29
 ; PRIOR APPLICATION NUMBER: 60/185,880
 ; PRIOR FILING DATE: 2000-02-29
 ; PRIOR APPLICATION NUMBER: 60/186,150
 ; PRIOR FILING DATE: 2000-03-01
 ; PRIOR APPLICATION NUMBER: 60/189,701
 ; PRIOR FILING DATE: 2000-03-15
 ; NUMBER OF SEQ ID NOS: 14
 ; SOFTWARE: PatentIn version 3.0
 ; SEQ ID NO 6
 ; LENGTH: 478
 ; TYPE: PRT
 ; ORGANISM: Drosophila melanogaster
 US-09-738-946-6

Query Match 56.2%; Score 54; DB 2; Length 478;
 Best Local Similarity 61.1%; Pred. No. 1.1;
 Matches 11; Conservative 3; Mismatches 4; Indels 0; Gaps 0;

QY 1 QDSKNNLTAQQQAQAQ 18
 DB 415 ENSKTNLAAQAQAQAE 432

RESULT 8
 US-09-134-000C-5990
 ; Sequence 5990, Application US/09134000C
 ; Patent No. 6617156
 ; GENERAL INFORMATION:
 ; APPLICANT: Lynn Doucette-Stamm et al
 ; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO
 ; FILE REFERENCE: 032796-032
 ; CURRENT APPLICATION NUMBER: US/09/134,000C
 ; CURRENT FILING DATE: 1998-08-13
 ; PRIOR APPLICATION NUMBER: US 60/055,778
 ; PRIOR FILING DATE: 1997-08-15
 ; NUMBER OF SEQ ID NOS: 6812
 ; SOFTWARE: PatentIn version 3.1
 ; SEQ ID NO 5990
 ; LENGTH: 497
 ; TYPE: PRT
 ; ORGANISM: Enterococcus faecalis
 US-09-134-000C-5990

Query Match 54.2%; Score 52; DB 2; Length 497;
 Best Local Similarity 57.9%; Pred. No. 2.4;
 Matches 11; Conservative 3; Mismatches 5; Indels 0; Gaps 0;

QY 1 QDSKNNLTAQQQAQAQ 19
 DB 52 QDOKINALTSQMSDAEAKV 70

RESULT 9
 US-09-550-163-4
 ; Sequence 4, Application US/09550163
 ; Patent No. 6864364
 ; GENERAL INFORMATION:
 ; APPLICANT: Abbott, Geoffrey W.
 ; APPLICANT: Sesti, Federico
 ; APPLICANT: Splawski, Igor
 ; APPLICANT: Keating, Mark T.
 ; APPLICANT: Goldstein, Steve A.N.
 ; APPLICANT: University of Utah Research Foundation
 ; APPLICANT: Yale University
 ; TITLE OF INVENTION: Mink-Related Genes, Formation of Potassium Channels and
 ; FILE REFERENCE: KCNE2 et al.
 ; CURRENT APPLICATION NUMBER: US/09/550,163
 ; CURRENT FILING DATE: 2000-04-14
 ; PRIOR APPLICATION NUMBER: US 60/129,404
 ; PRIOR FILING DATE: 1999-04-15
 ; NUMBER OF SEQ ID NOS: 20
 ; SOFTWARE: PatentIn Ver. 2.0
 ; SEQ ID NO 4
 ; LENGTH: 123
 ; TYPE: PRT
 ; ORGANISM: Rattus norvegicus
 US-09-550-163-4

Query Match 50.0%; Score 48; DB 2; Length 123;
 Best Local Similarity 57.9%; Pred. No. 2.1;
 Matches 11; Conservative 3; Mismatches 5; Indels 0; Gaps 0;

QY 2 DSKNNLTAQQQAQAQVN 20
 DB 24 DSWRRNTTAEQALQARVD 42

RESULT 10
 US-09-107-532A-5095
 ; Sequence 5095, Application US/09107532A
 ; Patent No. 6583275
 ; GENERAL INFORMATION:
 ; APPLICANT: Lynn A Doucette-Stamm and David Bush
 ; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO
 ; ENTEROCOCCUS FAECIUM FOR DIAGNOSTICS AND THERAPEUTICS
 ; NUMBER OF SEQUENCES: 7310
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: GENOME THERAPEUTICS CORPORATION
 ; STREET: 100 Beaver Street
 ; CITY: Waltham
 ; STATE: Massachusetts
 ; COUNTRY: USA
 ; ZIP: 02354
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: CD/ROM ISO9660
 ; COMPUTER: PC
 ; OPERATING SYSTEM: <Unknown>
 ; SOFTWARE: ASCII
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/09/107,532A
 ; FILING DATE: 30-Jun-1998
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER: 60/085,598
 ; FILING DATE: 14 May 1998
 ; APPLICATION NUMBER: 60/051571

MOLECULE LIFE:	/desc = protein fragment
DESCRIPTION:	

```

; ORGANISM: Mus musculus
US-09-978-309A-73

Query Match      46.9%; Score 45; DB 2; Length 333;
Best Local Similarity 44.4%; Pred. No. 20;
Matches      8; Conservative      5; Mismatches      5; Indels      0; Gaps      0;

QY      1 QDSKNNLTAQQQAQAAQ 18
Db       :| |: || |::||:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:~
121 EDLKLENTLTQEKVAMAE 138

RESULT 14
US-09-978-309A-80
; Sequence 80, Application US/09978309A
; Patent No. 6911429
; GENERAL INFORMATION:
; APPLICANT: Cruz, Tony
; APPLICANT: Pastrak, Aleksandra
; APPLICANT: Turley, Eva A.
; TITLE OF INVENTION: Compositions and Methods for Treating Cellular Response to Injury and Other Proliferating Cell Disorders Regulated by Hyaladherin and Hyaluronans
; TITLE OF INVENTION: Injury and Other Proliferating Cell Disorders Regulated by Hyaladherin and Hyaluronans
; FILE REFERENCE: 033352-010
; CURRENT APPLICATION NUMBER: US/09/978,309A
; CURRENT FILING DATE: 2002-04-04
; PRIOR APPLICATION NUMBER: US 09/685,010
; PRIOR FILING DATE: 2000-10-05
; PRIOR APPLICATION NUMBER: US 09/541,522
; PRIOR FILING DATE: 2000-04-03
; PRIOR APPLICATION NUMBER: US 60/127,457
; PRIOR FILING DATE: 1999-04-01
; NUMBER OF SEQ ID NOS: 84
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 80
; LENGTH: 435
; TYPE: PRT
; ORGANISM: Mus musculus
US-09-978-309A-80

Query Match      46.9%; Score 45; DB 2; Length 435;
Best Local Similarity 44.4%; Pred. No. 27;
Matches      8; Conservative      5; Mismatches      5; Indels      0; Gaps      0;

QY      1 QDSKNNLTAQQQAQAAQ 18
Db       :| |: || |::||:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:~
238 EDLKLENTLTQEKVAMAE 255

RESULT 15
US-09-978-309A-77
; Sequence 77, Application US/09978309A
; Patent No. 6911429
; GENERAL INFORMATION:
; APPLICANT: Cruz, Tony
; APPLICANT: Pastrak, Aleksandra
; APPLICANT: Turley, Eva A.
; TITLE OF INVENTION: Compositions and Methods for Treating Cellular Response to Injury and Other Proliferating Cell Disorders Regulated by Hyaladherin and Hyaluronans
; TITLE OF INVENTION: Injury and Other Proliferating Cell Disorders Regulated by Hyaladherin and Hyaluronans
; FILE REFERENCE: 033352-010
; CURRENT APPLICATION NUMBER: US/09/978,309A
; CURRENT FILING DATE: 2002-04-04
; PRIOR APPLICATION NUMBER: US 09/685,010
; PRIOR FILING DATE: 2000-10-05
; PRIOR APPLICATION NUMBER: US 09/541,522
; PRIOR FILING DATE: 2000-04-03
; PRIOR APPLICATION NUMBER: US 60/127,457
; PRIOR FILING DATE: 1999-04-01
; NUMBER OF SEQ ID NOS: 84
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 77
; LENGTH: 476

```

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2006 Compugen Ltd.

OM protein - protein search, using sw model

Run on: January 25, 2006, 19:13:51 ; Search time 60 Seconds
(without alignments)
139.276 Million cell updates/sec

Title: US-10-797-821-7

Perfect score: 96

Sequence: 1 QDSKINNLTAAQQAAQAQVN 20

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1867569 seqs, 417829326 residues

Total number of hits satisfying chosen parameters: 1867569

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications AA Main:

- 1: /cgn2_6/ptodata/1/pubpaa/US07_PUBCOMB.pep.*
- 2: /cgn2_6/ptodata/1/pubpaa/US08_PUBCOMB.pep.*
- 3: /cgn2_6/ptodata/1/pubpaa/US09_PUBCOMB.pep.*
- 4: /cgn2_6/ptodata/1/pubpaa/US10A_PUBCOMB.pep.*
- 5: /cgn2_6/ptodata/1/pubpaa/US10B_PUBCOMB.pep.*
- 6: /cgn2_6/ptodata/1/pubpaa/US11_PUBCOMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	96	100.0	20	4	US-10-383-930-7
2	96	100.0	20	5	US-10-797-821-7
3	96	100.0	431	4	US-10-383-930-29
4	96	100.0	431	4	US-10-383-930-30
5	96	100.0	431	4	US-10-383-930-33
6	96	100.0	431	5	US-10-797-821-29
7	96	100.0	431	5	US-10-797-821-30
8	96	100.0	431	5	US-10-797-821-33
9	96	100.0	432	4	US-10-383-930-31
10	96	100.0	432	4	US-10-383-930-32
11	96	100.0	432	5	US-10-797-821-31
12	96	100.0	432	5	US-10-797-821-32
13	77	80.2	20	4	US-10-383-930-6
14	77	80.2	20	5	US-10-797-821-6
15	73	76.0	210	4	US-10-154-251-67
16	73	76.0	392	5	US-10-472-928-4652
17	73	76.0	399	5	US-10-617-320-3230
18	61	63.5	398	5	US-10-474-792-600
19	54	56.2	325	6	US-11-097-143-117
20	52	54.2	258	4	US-10-424-599-239911
21	52	54.2	308	4	US-10-424-599-239913
22	50.5	52.6	1188	4	US-10-282-122A-71792
23	49	51.0	509	4	US-10-424-599-270324
24	48	50.0	123	4	US-10-842-558-4
25	48	50.0	618	6	US-11-097-143-23781
26	47	49.0	2238	3	US-09-980-217-20
27	47	49.0	2238	5	US-10-732-923-20547

```

28 46 47.9 524 4 US-10-282-122A-57658 Sequence 57658, A
29 45 46.9 190 4 US-10-767-701-32717 Sequence 32717, A
30 45 46.9 221 3 US-09-978-309A-75 Sequence 75, Appl
31 45 46.9 221 5 US-10-892-831-75 Sequence 75, Appl
32 45 46.9 221 5 US-10-994-157-3 Sequence 3, Appli
33 45 46.9 243 5 US-10-805-684-18 Sequence 18, Appl
34 45 46.9 302 4 US-10-108-260A-2659 Sequence 2659, Ap
35 45 46.9 333 3 US-09-978-309A-73 Sequence 73, Appl
36 45 46.9 333 5 US-10-892-831-73 Sequence 73, Appl
37 45 46.9 333 5 US-10-994-157-1 Sequence 1, Appli
38 45 46.9 435 3 US-09-978-309A-80 Sequence 80, Appl
39 45 46.9 435 5 US-10-892-831-80 Sequence 80, Appl
40 45 46.9 435 5 US-10-994-157-10 Sequence 10, Appl
41 45 46.9 476 3 US-09-978-309A-77 Sequence 77, Appl
42 45 46.9 476 3 US-09-978-309A-79 Sequence 79, Appl
43 45 46.9 476 5 US-10-892-831-77 Sequence 77, Appl
44 45 46.9 476 5 US-10-892-831-79 Sequence 79, Appl
45 45 46.9 477 5 US-10-994-157-7 Sequence 7, Appli

```

ALIGNMENTS

RESULT 1

```

US-10-383-930-7
; Sequence 7, Application US/10383930
; Publication No. US20040127400A1
; GENERAL INFORMATION:
; APPLICANT: Smith, Daniel J
; APPLICANT: Taubman, Martin A
; TITLE OF INVENTION: Immunogenicity of Glucan Binding Protein
; FILE REFERENCE: 25669-018
; CURRENT APPLICATION NUMBER: US/10/383,930
; CURRENT FILING DATE: 2003-03-07
; PRIOR APPLICATION NUMBER: 60/402,483
; PRIOR FILING DATE: 2002-08-08
; PRIOR APPLICATION NUMBER: 60/363,209
; PRIOR FILING DATE: 2002-03-07
; NUMBER OF SEQ ID NOS: 41
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 7
; LENGTH: 20
; TYPE: PRT
; ORGANISM: Streptococcus mutans
US-10-383-930-7

```

```

Query Match 100.0%; Score 96; DB 4; Length 20;
Best Local Similarity 100.0%; Pred. No. 3.7e-08;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

Qy 1 QDSKINNLTAAQQAAQAQVN 20
      |||||
Db 1 QDSKINNLTAAQQAAQAQVN 20

```

RESULT 2

```

US-10-797-821-7
; Sequence 7, Application US/10797821
; Publication No. US20050031633A1
; GENERAL INFORMATION:
; APPLICANT: Smith, Daniel J.
; APPLICANT: Taubman, Martin A.
; TITLE OF INVENTION: Glucan Binding Protein and Glycosyltransferase Immunogens
; FILE REFERENCE: 25669-020
; CURRENT APPLICATION NUMBER: US/10/797,821
; CURRENT FILING DATE: 2004-03-09
; PRIOR APPLICATION NUMBER: 10/383,930
; PRIOR FILING DATE: 2003-03-07
; PRIOR APPLICATION NUMBER: 60/363,209
; PRIOR FILING DATE: 2002-03-07
; PRIOR APPLICATION NUMBER: 60/402,483
; PRIOR FILING DATE: 2002-08-08
; PRIOR APPLICATION NUMBER: 09/290,049

```

; PRIOR FILING DATE: 1999-04-12
; PRIOR APPLICATION NUMBER: 60/081,550
; PRIOR FILING DATE: 1998-04-13
; PRIOR APPLICATION NUMBER: 60/115,142
; PRIOR FILING DATE: 1999-01-08
; NUMBER OF SEQ ID NOS: 45
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 7
; LENGTH: 20
; TYPE: PRT
; ORGANISM: Artificial
; FEATURE:
; OTHER INFORMATION: GbpB peptide
US-10-797-821-7

Query Match 100.0%; Score 96; DB 5; Length 20;
Best Local Similarity 100.0%; Pred. No. 3.7e-08;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 QDSKINNLTAAQQAAQAQVN 20
Db 1 QDSKINNLTAAQQAAQAQVN 20
|||||

RESULT 3
US-10-383-930-29
; Sequence 29, Application US/10383930
; Publication No. US20040127400A1
; GENERAL INFORMATION:
; APPLICANT: Smith, Daniel J
; APPLICANT: Taubman, Martin A
; TITLE OF INVENTION: Immunogenicity of Glucan Binding Protein
; FILE REFERENCE: 25669-018
; CURRENT APPLICATION NUMBER: US/10/383,930
; CURRENT FILING DATE: 2003-03-07
; PRIOR APPLICATION NUMBER: 60/402,483
; PRIOR FILING DATE: 2002-08-08
; PRIOR APPLICATION NUMBER: 60/363,209
; PRIOR FILING DATE: 2002-03-07
; NUMBER OF SEQ ID NOS: 41
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 29
; LENGTH: 431
; TYPE: PRT
; ORGANISM: Streptococcus mutans
US-10-383-930-29

Query Match 100.0%; Score 96; DB 4; Length 431;
Best Local Similarity 100.0%; Pred. No. 1.1e-06;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 QDSKINNLTAAQQAAQAQVN 20
Db 37 QDSKINNLTAAQQAAQAQVN 56
|||||

RESULT 4
US-10-383-930-30
; Sequence 30, Application US/10383930
; Publication No. US20040127400A1
; GENERAL INFORMATION:
; APPLICANT: Smith, Daniel J
; APPLICANT: Taubman, Martin A
; TITLE OF INVENTION: Immunogenicity of Glucan Binding Protein
; FILE REFERENCE: 25669-018
; CURRENT APPLICATION NUMBER: US/10/383,930
; CURRENT FILING DATE: 2003-03-07
; PRIOR APPLICATION NUMBER: 60/402,483
; PRIOR FILING DATE: 2002-08-08
; PRIOR APPLICATION NUMBER: 60/363,209
; PRIOR FILING DATE: 2002-03-07
; NUMBER OF SEQ ID NOS: 41
; SOFTWARE: PatentIn version 3.2

; SEQ ID NO 30
; LENGTH: 431
; TYPE: PRT
; ORGANISM: Streptococcus mutans
US-10-383-930-30

Query Match 100.0%; Score 96; DB 4; Length 431;
Best Local Similarity 100.0%; Pred. No. 1.1e-06;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 QDSKINNLTAAQQAAQAQVN 20
Db 37 QDSKINNLTAAQQAAQAQVN 56
|||||

RESULT 5
US-10-383-930-33
; Sequence 33, Application US/10383930
; Publication No. US20040127400A1
; GENERAL INFORMATION:
; APPLICANT: Smith, Daniel J
; APPLICANT: Taubman, Martin A
; TITLE OF INVENTION: Immunogenicity of Glucan Binding Protein
; FILE REFERENCE: 25669-018
; CURRENT APPLICATION NUMBER: US/10/383,930
; CURRENT FILING DATE: 2003-03-07
; PRIOR APPLICATION NUMBER: 60/402,483
; PRIOR FILING DATE: 2002-08-08
; PRIOR APPLICATION NUMBER: 60/363,209
; PRIOR FILING DATE: 2002-03-07
; NUMBER OF SEQ ID NOS: 41
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 33
; LENGTH: 431
; TYPE: PRT
; ORGANISM: Streptococcus mutans
US-10-383-930-33

Query Match 100.0%; Score 96; DB 4; Length 431;
Best Local Similarity 100.0%; Pred. No. 1.1e-06;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 QDSKINNLTAAQQAAQAQVN 20
Db 37 QDSKINNLTAAQQAAQAQVN 56
|||||

RESULT 6
US-10-797-821-29
; Sequence 29, Application US/10797821
; Publication No. US20050031633A1
; GENERAL INFORMATION:
; APPLICANT: Smith, Daniel J.
; APPLICANT: Taubman, Martin A.
; TITLE OF INVENTION: Glucan Binding Protein and Glycosyltransferase Immunogens
; FILE REFERENCE: 25669-020
; CURRENT APPLICATION NUMBER: US/10/797,821
; CURRENT FILING DATE: 2004-03-09
; PRIOR APPLICATION NUMBER: 10/383,930
; PRIOR FILING DATE: 2003-03-07
; PRIOR APPLICATION NUMBER: 60/363,209
; PRIOR FILING DATE: 2002-03-07
; PRIOR APPLICATION NUMBER: 60/402,483
; PRIOR FILING DATE: 2002-08-08
; PRIOR APPLICATION NUMBER: 09/290,049
; PRIOR FILING DATE: 1999-04-12
; PRIOR APPLICATION NUMBER: 60/081,550
; PRIOR FILING DATE: 1998-04-13
; PRIOR APPLICATION NUMBER: 60/115,142
; PRIOR FILING DATE: 1999-01-08
; NUMBER OF SEQ ID NOS: 45
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 29

```
; LENGTH: 431
; TYPE: PRT
; ORGANISM: Streptococcus mutans
US-10-797-821-29

Query Match      100.0%; Score 96; DB 5; Length 431;
Best Local Similarity 100.0%; Pred. No. 1.1e-06;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 QDSKINNLTAAQQAQAQVN 20
   |||||
Db 37 QDSKINNLTAAQQAQAQVN 56

RESULT 7
US-10-797-821-30
; Sequence 30, Application US/10797821
; Publication No. US20050031633A1
; GENERAL INFORMATION:
; APPLICANT: Smith, Daniel J.
; APPLICANT: Taubman, Martin A.
; TITLE OF INVENTION: Glucan Binding Protein and Glycosyltransferase Immunogens
; FILE REFERENCE: 25669-020
; CURRENT APPLICATION NUMBER: US/10/797,821
; CURRENT FILING DATE: 2004-03-09
; PRIOR APPLICATION NUMBER: 10/383,930
; PRIOR FILING DATE: 2003-03-07
; PRIOR APPLICATION NUMBER: 60/363,209
; PRIOR FILING DATE: 2002-03-07
; PRIOR APPLICATION NUMBER: 60/402,483
; PRIOR FILING DATE: 2002-08-08
; PRIOR APPLICATION NUMBER: 09/290,049
; PRIOR FILING DATE: 1999-04-12
; PRIOR APPLICATION NUMBER: 60/081,550
; PRIOR FILING DATE: 1998-04-13
; PRIOR APPLICATION NUMBER: 60/115,142
; PRIOR FILING DATE: 1999-01-08
; NUMBER OF SEQ ID NOS: 45
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 30
; LENGTH: 431
; TYPE: PRT
; ORGANISM: Streptococcus mutans
US-10-797-821-30

Query Match      100.0%; Score 96; DB 5; Length 431;
Best Local Similarity 100.0%; Pred. No. 1.1e-06;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 QDSKINNLTAAQQAQAQVN 20
   |||||
Db 37 QDSKINNLTAAQQAQAQVN 56

RESULT 8
US-10-797-821-33
; Sequence 33, Application US/10797821
; Publication No. US20050031633A1
; GENERAL INFORMATION:
; APPLICANT: Smith, Daniel J.
; APPLICANT: Taubman, Martin A.
; TITLE OF INVENTION: Glucan Binding Protein and Glycosyltransferase Immunogens
; FILE REFERENCE: 25669-020
; CURRENT APPLICATION NUMBER: US/10/797,821
; CURRENT FILING DATE: 2004-03-09
; PRIOR APPLICATION NUMBER: 10/383,930
; PRIOR FILING DATE: 2003-03-07
; PRIOR APPLICATION NUMBER: 60/363,209
; PRIOR FILING DATE: 2002-03-07
; PRIOR APPLICATION NUMBER: 60/402,483
; PRIOR FILING DATE: 2002-08-08
; PRIOR APPLICATION NUMBER: 09/290,049
; PRIOR FILING DATE: 1999-04-12
; NUMBER OF SEQ ID NOS: 45
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 30
; LENGTH: 431
; TYPE: PRT
; ORGANISM: Streptococcus mutans
US-10-797-821-33

Query Match      100.0%; Score 96; DB 5; Length 431;
Best Local Similarity 100.0%; Pred. No. 1.1e-06;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 QDSKINNLTAAQQAQAQVN 20
   |||||
Db 37 QDSKINNLTAAQQAQAQVN 56

RESULT 9
US-10-383-930-31
; Sequence 31, Application US/10383930
; Publication No. US20040127400A1
; GENERAL INFORMATION:
; APPLICANT: Smith, Daniel J.
; APPLICANT: Taubman, Martin A.
; TITLE OF INVENTION: Immunogenicity of Glucan Binding Protein
; FILE REFERENCE: 25669-018
; CURRENT APPLICATION NUMBER: US/10/383,930
; CURRENT FILING DATE: 2003-03-07
; PRIOR APPLICATION NUMBER: 60/402,483
; PRIOR FILING DATE: 2002-08-08
; PRIOR APPLICATION NUMBER: 60/363,209
; PRIOR FILING DATE: 2002-03-07
; NUMBER OF SEQ ID NOS: 41
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 31
; LENGTH: 432
; TYPE: PRT
; ORGANISM: Streptococcus mutans
US-10-383-930-31

Query Match      100.0%; Score 96; DB 4; Length 432;
Best Local Similarity 100.0%; Pred. No. 1.1e-06;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 QDSKINNLTAAQQAQAQVN 20
   |||||
Db 37 QDSKINNLTAAQQAQAQVN 56

RESULT 10
US-10-383-930-32
; Sequence 32, Application US/10383930
; Publication No. US20040127400A1
; GENERAL INFORMATION:
; APPLICANT: Smith, Daniel J.
; APPLICANT: Taubman, Martin A.
; TITLE OF INVENTION: Immunogenicity of Glucan Binding Protein
; FILE REFERENCE: 25669-018
; CURRENT APPLICATION NUMBER: US/10/383,930
; CURRENT FILING DATE: 2003-03-07
; PRIOR APPLICATION NUMBER: 60/402,483
; PRIOR FILING DATE: 2002-08-08
; PRIOR APPLICATION NUMBER: 60/363,209
; PRIOR FILING DATE: 2002-03-07
; NUMBER OF SEQ ID NOS: 41
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 32
; LENGTH: 432
; TYPE: PRT
; ORGANISM: Streptococcus mutans
US-10-383-930-32
```

	; PRIOR APPLICATION NUMBER: 60/115,142	
	; PRIOR FILING DATE: 1999-01-08	
	; NUMBER OF SEQ ID NOS: 45	
	; SOFTWARE: PatentIn version 3.2	
	; SEQ ID NO 32	
	; LENGTH: 432	
	; TYPE: PRT	
	; ORGANISM: Streptococcus mutans	
	US-10-797-821-32	
	Query Match 100.0%; Score 96; DB 5; Length 432;	
	Best Local Similarity 100.0%; Pred. No. 1.1e-06;	
	Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;	
QY	1 QDSKNNLTAQQAAQAQYN 20 	
DB	37 QDSKNNLTAQQAAQAQYN 56 	
RESULT 13		
US-10-383-930-6		
	; Sequence 6, Application US/10383930	
	; Publication No. US20040127400A1	
	; GENERAL INFORMATION:	
	; APPLICANT: Smith, Daniel J.	
	; APPLICANT: Taubman, Martin A.	
	; TITLE OF INVENTION: Immunogenicity of Glucan Binding Protein	
	; FILE REFERENCE: 25669-018	
	; CURRENT APPLICATION NUMBER: US/10/383,930	
	; CURRENT FILING DATE: 2003-03-07	
	; PRIOR APPLICATION NUMBER: 60/402,483	
	; PRIOR FILING DATE: 2002-08-08	
	; PRIOR APPLICATION NUMBER: 60/363,209	
	; PRIOR FILING DATE: 2002-03-07	
	; NUMBER OF SEQ ID NOS: 41	
	; SOFTWARE: PatentIn version 3.2	
	; SEQ ID NO 6	
	; LENGTH: 20	
	; TYPE: PRT	
	; ORGANISM: Streptococcus mutans	
	US-10-383-930-6	
	Query Match 80.2%; Score 77; DB 4; Length 20;	
	Best Local Similarity 100.0%; Pred. No. 3.6e-05;	
	Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;	
QY	1 QDSKNNLTAQQAAAQ 16 	
DB	5 QDSKNNLTAQQAAAQ 20 	
RESULT 14		
US-10-797-821-6		
	; Sequence 6, Application US/10797821	
	; Publication No. US20050031633A1	
	; GENERAL INFORMATION:	
	; APPLICANT: Smith, Daniel J.	
	; APPLICANT: Taubman, Martin A.	
	; TITLE OF INVENTION: Glucan Binding Protein and Glycosyltransferase Immunogen	
	; FILE REFERENCE: 25669-020	
	; CURRENT APPLICATION NUMBER: US/10/797,821	
	; CURRENT FILING DATE: 2004-03-09	
	; PRIOR APPLICATION NUMBER: 10/383,930	
	; PRIOR FILING DATE: 2003-03-07	
	; PRIOR APPLICATION NUMBER: 60/363,209	
	; PRIOR FILING DATE: 2002-03-07	
	; PRIOR APPLICATION NUMBER: 60/402,483	
	; PRIOR FILING DATE: 2002-08-08	
	; PRIOR APPLICATION NUMBER: 09/290,049	
	; PRIOR FILING DATE: 1999-04-12	
	; PRIOR APPLICATION NUMBER: 60/081,550	
	; PRIOR FILING DATE: 1998-04-13	
	; PRIOR APPLICATION NUMBER: 60/115,142	

; PRIOR FILING DATE: 1999-01-08
; NUMBER OF SEQ ID NOS: 45
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 6
; LENGTH: 20
; TYPE: PRT
; ORGANISM: Artificial
; FEATURE:
; OTHER INFORMATION: GbpB peptide
US-10-797-821-6

Query Match 80.2%; Score 77; DB 5; Length 20;
Best Local Similarity 100.0%; Pred. No. 3.6e-05;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 QDSKINNLTAAQQAAQ 16
Db 5 QDSKINNLTAAQQAAQ 20

RESULT 15
US-10-154-251-67
; Sequence 67, Application US/10154251
; Publication No. US20030092024A1
; GENERAL INFORMATION:
; APPLICANT: Youngman, Philip
; APPLICANT: Fritz, Christian
; APPLICANT: Murphy, Christopher
; APPLICANT: Guzman, Luz-Maria
; TITLE OF INVENTION: ESSENTIAL BACTERIAL GENES AND THEIR USE
; FILE REFERENCE: 06286-060002
; CURRENT APPLICATION NUMBER: US/10/154,251
; CURRENT FILING DATE: 2002-09-16
; PRIOR APPLICATION NUMBER: US 10/154,251
; PRIOR FILING DATE: 2002-05-22
; NUMBER OF SEQ ID NOS: 102
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 67
; LENGTH: 210
; TYPE: PRT
; ORGANISM: Streptococcus pneumoniae
US-10-154-251-67

Query Match 76.0%; Score 73; DB 4; Length 210;
Best Local Similarity 75.0%; Pred. No. 0.0021;
Matches 15; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

Qy 1 QDSKINNLTAAQQAAQVN 20
Db 37 QDNKISNLTAAQQEAQRQVD 56

Search completed: January 25, 2006, 20:10:03
Job time : 60 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2006 Compugen Ltd.

OM protein - protein search, using sw model

Run on: January 25, 2006, 19:15:26 (Search time 5.5 Seconds
(without alignments))
39.378 Million cell updates/sec

Title: US-10-797-821-7

Perfect score: 96

Sequence: 1 QDSKNNLTAAQQAAQAVN 20

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 75621 seqs, 10829074 residues

Total number of hits satisfying chosen parameters: 75621

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications AA New:*

- 1: /cgn2_6/ptodata/2/pubpaa/US08_NEW_PUB.pep.*
- 2: /cgn2_6/ptodata/2/pubpaa/US06_NEW_PUB.pep.*
- 3: /cgn2_6/ptodata/2/pubpaa/US07_NEW_PUB.pep.*
- 4: /cgn2_6/ptodata/2/pubpaa/ECT_NEW_PUB.pep.*
- 5: /cgn2_6/ptodata/2/pubpaa/US09_NEW_PUB.pep.*
- 6: /cgn2_6/ptodata/2/pubpaa/US10_NEW_PUB.pep.*
- 7: /cgn2_6/ptodata/2/pubpaa/US11_NEW_PUB.pep.*
- 8: /cgn2_6/ptodata/2/pubpaa/US60_NEW_PUB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	96	100.0	431	7 US-11-052-554A-210	Sequence 210, App
2	61	63.5	398	7 US-11-052-554A-252	Sequence 252, App
3	45	46.9	205	6 US-10-467-657-4336	Sequence 433, App
4	44	45.8	515	7 US-11-132-142-8	Sequence 8, Appli
5	43.5	45.3	885	6 US-10-793-626-1660	Sequence 166, Ap
6	40	41.7	405	6 US-10-467-657-3078	Sequence 307, Ap
7	40	41.7	716	7 US-11-194-890-20	Sequence 20, Appl
8	39.5	41.1	706	6 US-10-485-517-146	Sequence 146, App
9	39	40.6	298	6 US-10-467-657-2850	Sequence 285, App
10	39	40.6	298	6 US-10-467-657-750	Sequence 675, Ap
11	39	40.6	612	6 US-10-467-657-3988	Sequence 398, Ap
12	38	39.6	1107	6 US-10-485-517-145	Sequence 145, App
13	37	38.5	1279	6 US-10-793-626-3188	Sequence 318, App
14	37	38.5	1960	7 US-11-069-834-48	Sequence 48, Appl
15	36.5	38.0	1095	6 US-10-793-626-3154	Sequence 315, Ap
16	36	37.5	616	7 US-11-083-800-4	Sequence 4, Appli
17	36	37.5	565	6 US-10-873-528-127	Sequence 127, App
18	36	37.5	771	6 US-10-467-657-5562	Sequence 556, Ap
19	36	37.5	860	7 US-11-019-711-59	Sequence 59, Appl
20	36	37.5	1242	6 US-10-517-939-308	Sequence 308, App
21	36	37.5	1290	6 US-10-485-517-141	Sequence 141, App
22	35	36.5	302	6 US-10-793-626-3062	Sequence 306, Ap
23	35	36.5	352	6 US-10-793-626-216	Sequence 216, App
24	35	36.5	507	7 US-11-052-554A-216	Sequence 216, App
25	35	36.5	627	6 US-10-467-657-5432	Sequence 543, Ap

ALIGNMENTS

RESULT 1

US-11-052-554A-210
; Sequence 210, Application US/11052554A
; Publication No. US20050288866A1
; GENERAL INFORMATION:

; APPLICANT: Sachdeva, et al.
; TITLE OF INVENTION: COMPUTATIONAL METHOD FOR IDENTIFYING ADHESIN AND ADHESIN-LIKE
; TITLE OF INVENTION: PROTEINS OF THERAPEUTIC POTENTIAL
; FILE REFERENCE: 30853/40359A

; CURRENT APPLICATION NUMBER: US/11/052,554A

; CURRENT FILING DATE: 2005-02-07

; PRIOR APPLICATION NUMBER: US 60/589,227

; PRIOR FILING DATE: 2004-07-20

; PRIOR APPLICATION NUMBER: IN 173/DEL/2004

; PRIOR FILING DATE: 2004-02-06

; NUMBER OF SEQ ID NOS: 763

; SOFTWARE: PatentIn version 3.3

; SEQ ID NO 210

; LENGTH: 431

; TYPE: PRT

; ORGANISM: Streptococcus mutans UA159

US-11-052-554A-210

Query Match 100.0%; Score 96; DB 7; Length 431;

Best Local Similarity 100.0%; Pred. No. 2e-08;

Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 QDSKNNLTAAQQAAQAVN 20

|||||

Db 37 QDSKNNLTAAQQAAQAVN 56

|||||

RESULT 2

US-11-052-554A-252

; Sequence 252, Application US/11052554A

; Publication No. US20050288866A1

; GENERAL INFORMATION:

; APPLICANT: Sachdeva, et al.

; TITLE OF INVENTION: COMPUTATIONAL METHOD FOR IDENTIFYING ADHESIN AND ADHESIN-LIKE
; TITLE OF INVENTION: PROTEINS OF THERAPEUTIC POTENTIAL

; FILE REFERENCE: 30853/40359A

; CURRENT APPLICATION NUMBER: US/11/052,554A

; CURRENT FILING DATE: 2005-02-07

; PRIOR APPLICATION NUMBER: US 60/589,227

; PRIOR FILING DATE: 2004-07-20

; PRIOR APPLICATION NUMBER: IN 173/DEL/2004

; PRIOR FILING DATE: 2004-02-06

; NUMBER OF SEQ ID NOS: 763

Sequence 2484, Ap
Sequence 1002, Ap
Sequence 676, App
Sequence 20, Appl
Sequence 108, App
Sequence 1514, Ap
Sequence 1003, Ap
Sequence 1000, Ap
Sequence 1004, Ap
Sequence 1001, Ap
Sequence 998, App
Sequence 997, App
Sequence 7930, Ap
Sequence 52, Appl
Sequence 53, Appl
Sequence 54, Appl
Sequence 55, Appl
Sequence 212, App
Sequence 48, Appl
Sequence 51, Appl

```
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 252
; LENGTH: 398
; TYPE: PRT
; ORGANISM: Streptococcus pyogenes MGAS8232
US-11-052-554A-252

Query Match      63.5%; Score 61; DB 7; Length 398;
Best Local Similarity 65.0%; Pred. No. 0.0086;
Matches 13; Conservative 4; Mismatches 3; Indels 0; Gaps 0;

Qy 1 QDSKINNLTAQQAQAQVN 20
Db 34 QDSIISNLTTEQAQAQNV 53

RESULT 3
US-10-467-657-4336
; Sequence 4336, Application US/10467657
; Publication No. US20050260581A1
; GENERAL INFORMATION:
; APPLICANT: CHIRON Spa
; APPLICANT: FONTANA Maria Rita
; APPLICANT: PIZZA Mariagrazia
; APPLICANT: MASNANI Vega
; APPLICANT: MONACI Elisabetta
; TITLE OF INVENTION: GONOCOCCAL PROTEINS AND NUCLEIC ACIDS
; FILE REFERENCE:
; CURRENT APPLICATION NUMBER: US/10/467,657
; CURRENT FILING DATE: 2003-08-11
; PRIOR APPLICATION NUMBER: GB-0103424.8
; PRIOR FILING DATE: 2001-02-12
; NUMBER OF SEQ ID NOS: 9218
; SOFTWARE: SeqWin99, version 1.04
; SEQ ID NO 4336
; LENGTH: 205
; TYPE: PRT
; ORGANISM: Neisseria gonorrhoeae
US-10-467-657-4336

Query Match      46.9%; Score 45; DB 6; Length 205;
Best Local Similarity 55.6%; Pred. No. 1.6;
Matches 10; Conservative 3; Mismatches 5; Indels 0; Gaps 0;

Qy 1 QDSKINNLTAQQAQAQ 18
Db 164 RDLNTRSLRAKQAQAQ 181

RESULT 4
US-11-132-142-8
; Sequence 8, Application US/11132142
; Publication No. US2005027681A1
; GENERAL INFORMATION:
; APPLICANT: The Burnham Institute
; APPLICANT: Sikora, Sergey
; APPLICANT: Godzik, Adam
; TITLE OF INVENTION: UNCHARACTERIZED ORF3 IN SARS-CORONAVIRUS IS A
; FILE REFERENCE: CYCLIC-AMP-DEPENDENT KINASE AND A TARGET FOR SARS THERAPY
; CURRENT APPLICATION NUMBER: US/11/132,142
; CURRENT FILING DATE: 2005-05-17
; PRIOR APPLICATION NUMBER: 60/571,698
; PRIOR FILING DATE: 2004-05-17
; NUMBER OF SEQ ID NOS: 21
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 8
; LENGTH: 515
; TYPE: PRT
; ORGANISM: Cryptococcus sp.
US-11-132-142-8

Query Match      45.8%; Score 44; DB 7; Length 515;
Best Local Similarity 50.0%; Pred. No. 22;
Matches 9; Conservative 4; Mismatches 5; Indels 0; Gaps 0;

; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 252
; LENGTH: 398
; TYPE: PRT
; ORGANISM: Streptococcus pyogenes MGAS8232
US-10-797-821-7.rapbn

Query Match      81.8%; Pred. No. 6.5;
Matches 9; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 10 AQQAQAQVN 20
Db 134 AQQAQAQVN 144

RESULT 5
US-10-793-626-1660
; Sequence 1660, Application US/10793626
; Publication No. US20050255478A1
; GENERAL INFORMATION:
; APPLICANT: KIMMERLY, WILLIAM JOHN
; TITLE OF INVENTION: STAPHYLOCOCCUS EPIDERMIDIS NUCLEIC ACIDS AND PROTEINS
; FILE REFERENCE: P03480US
; CURRENT APPLICATION NUMBER: US/10/793,626
; CURRENT FILING DATE: 2004-03-04
; PRIOR APPLICATION NUMBER: 60/164,258
; PRIOR FILING DATE: 1999-11-09
; NUMBER OF SEQ ID NOS: 4472
; SOFTWARE: Patent Ver. 2.1
; SEQ ID NO 1660
; LENGTH: 885
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: synthetic
US-10-793-626-1660

Query Match      45.3%; Score 43.5; DB 6; Length 885;
Best Local Similarity 40.9%; Pred. No. 15;
Matches 9; Conservative 6; Mismatches 4; Indels 3; Gaps 1;

Qy 2 DSKINNLTAQQAQAQVN 20
Db 44 DERLNHLKSQAQKEGQAQIN 65

RESULT 6
US-10-467-657-3078
; Sequence 3078, Application US/10467657
; Publication No. US20050260581A1
; GENERAL INFORMATION:
; APPLICANT: CHIRON Spa
; APPLICANT: FONTANA Maria Rita
; APPLICANT: PIZZA Mariagrazia
; APPLICANT: MASNANI Vega
; APPLICANT: MONACI Elisabetta
; TITLE OF INVENTION: GONOCOCCAL PROTEINS AND NUCLEIC ACIDS
; FILE REFERENCE:
; CURRENT APPLICATION NUMBER: US/10/467,657
; CURRENT FILING DATE: 2003-08-11
; PRIOR APPLICATION NUMBER: GB-0103424.8
; PRIOR FILING DATE: 2001-02-12
; NUMBER OF SEQ ID NOS: 9218
; SOFTWARE: SeqWin99, version 1.04
; SEQ ID NO 3078
; LENGTH: 405
; TYPE: PRT
; ORGANISM: Neisseria gonorrhoeae
US-10-467-657-3078

Query Match      41.7%; Score 40; DB 6; Length 405;
Best Local Similarity 50.0%; Pred. No. 22;
Matches 9; Conservative 4; Mismatches 5; Indels 0; Gaps 0;

Qy 3 SKINNLTAQQAQAQVN 20
Db 371 AKVDETAQSQAKQAQVN 388
```

```

RESULT 7
US-11-194-890-20
; Sequence 20, Application US/11194890
; Publication No. US20050287579A1
; GENERAL INFORMATION:
; APPLICANT: Wolfner, Mariana
; APPLICANT: Lung, Oliver
; APPLICANT: Tram, Khanh-Uyen
; TITLE OF INVENTION: GENES FOR MALE ACCESSORY GLAND PROTEINS IN DROSOPHILA
; FILE REFERENCE: 19603/1791
; CURRENT APPLICATION NUMBER: US/11/194,890
; CURRENT FILING DATE: 2005-08-01
; PRIOR FILING DATE: 2002-04-02
; PRIOR APPLICATION NUMBER: 09/219,983
; PRIOR FILING DATE: 1998-12-23
; NUMBER OF SEQ ID NOS: 35
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 20
; LENGTH: 716
; TYPE: PRT
; ORGANISM: Drosophila melanogaster
US-11-194-890-20

Query Match 41.7%; Score 40; DB 7; Length 716;
Best Local Similarity 42.1%; Pred. No. 42;
Matches 8; Conservative 5; Mismatches 6; Indels 0; Gaps 0;

QY 1 QDSKNNLTAAQQQAQAV 19
Db 337 QESKNSLSQSQSQSQEQL 355

RESULT 8
US-10-485-517-146
; Sequence 146, Application US/10485517
; Publication No. US20050256299A1
; GENERAL INFORMATION:
; APPLICANT: University of Sheffield
; APPLICANT: Biosynexus Incorporated
; APPLICANT: Foster, Simon
; APPLICANT: Mond, James
; TITLE OF INVENTION: Antigenic Polypeptides
; FILE REFERENCE: P1006290
; CURRENT APPLICATION NUMBER: US/10/485,517
; CURRENT FILING DATE: 2004-02-02
; PRIOR FILING DATE: 2001-08-02
; PRIOR APPLICATION NUMBER: GB 0118825.9
; PRIOR FILING DATE: 2002-01-09
; NUMBER OF SEQ ID NOS: 424
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 146
; LENGTH: 706
; TYPE: PRT
; ORGANISM: Staphylococcus aureus
US-10-485-517-146

Query Match 41.1%; Score 39.5; DB 6; Length 706;
Best Local Similarity 52.9%; Pred. No. 50;
Matches 9; Conservative 4; Mismatches 3; Indels 1; Gaps 1;

QY 5 INNL-TAQQQAQAVN 20
Db 467 LNNLNAQRNLQSQIN 483

RESULT 9
US-10-467-657-2850
; Sequence 2850, Application US/10467657
; Publication No. US20050260581A1
; GENERAL INFORMATION:
; APPLICANT: CHIRON Spa
; APPLICANT: FONTANA Maria Rita
; APPLICANT: PIZZA Mariagrazia
; APPLICANT: MASIGNANI Vega
; TITLE OF INVENTION: GONOCOCCAL PROTEINS AND NUCLEIC ACIDS
; FILE REFERENCE:
; CURRENT APPLICATION NUMBER: US/10/467,657
; CURRENT FILING DATE: 2003-08-11
; PRIOR APPLICATION NUMBER: GB-0103424.8
; PRIOR FILING DATE: 2001-02-12
; NUMBER OF SEQ ID NOS: 9218
; SOFTWARE: SeqWin99, version 1.04
; SEQ ID NO 2850
; LENGTH: 298
; TYPE: PRT
; ORGANISM: Neisseria gonorrhoeae
US-10-467-657-2850

Query Match 40.6%; Score 39; DB 6; Length 298;
Best Local Similarity 69.2%; Pred. No. 22;
Matches 9; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 6 NNLTAAQQQAQAO 18
Db 5 NDLKAWQQAQAO 17

RESULT 10
US-10-467-657-6750
; Sequence 6750, Application US/10467657
; Publication No. US20050260581A1
; GENERAL INFORMATION:
; APPLICANT: CHIRON Spa
; APPLICANT: FONTANA Maria Rita
; APPLICANT: PIZZA Mariagrazia
; APPLICANT: MASIGNANI Vega
; APPLICANT: MONACI Elisabetta
; TITLE OF INVENTION: GONOCOCCAL PROTEINS AND NUCLEIC ACIDS
; FILE REFERENCE:
; CURRENT APPLICATION NUMBER: US/10/467,657
; CURRENT FILING DATE: 2003-08-11
; PRIOR APPLICATION NUMBER: GB-0103424.8
; PRIOR FILING DATE: 2001-02-12
; NUMBER OF SEQ ID NOS: 9218
; SOFTWARE: SeqWin99, version 1.04
; SEQ ID NO 6750
; LENGTH: 298
; TYPE: PRT
; ORGANISM: Neisseria gonorrhoeae
US-10-467-657-6750

Query Match 40.6%; Score 39; DB 6; Length 298;
Best Local Similarity 69.2%; Pred. No. 22;
Matches 9; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 6 NNLTAAQQQAQAO 18
Db 5 NDLKAWQQAQAO 17

RESULT 11
US-10-467-657-3988
; Sequence 3988, Application US/10467657
; Publication No. US20050260581A1
; GENERAL INFORMATION:
; APPLICANT: CHIRON Spa
; APPLICANT: FONTANA Maria Rita
; APPLICANT: PIZZA Mariagrazia
; APPLICANT: MASIGNANI Vega
; APPLICANT: MONACI Elisabetta
; TITLE OF INVENTION: GONOCOCCAL PROTEINS AND NUCLEIC ACIDS
; FILE REFERENCE:

```

```
; CURRENT APPLICATION NUMBER: US/10/467,657
; CURRENT FILING DATE: 2003-08-11
; PRIOR APPLICATION NUMBER: GB-0103424.8
; PRIOR FILING DATE: 2001-02-12
; NUMBER OF SEQ ID NOS: 9218
; SOFTWARE: SeqWin99, version 1.04
; SEQ ID NO 3988
; LENGTH: 612
; TYPE: PRT
; ORGANISM: Neisseria gonorrhoeae
US-10-467-657-3988

Query Match      40.6%; Score 39; DB 6; Length 612;
Best Local Similarity 47.4%; Pred. No. 51;
Matches 9; Conservative 3; Mismatches 7; Indels 0; Gaps 0;

Qy      1 QDSKNNLTAQQQAAQAV 19
Db      351 QEQKNNELARLKKIQNV 369

RESULT 12
US-10-485-517-145
; Sequence 145, Application US/10485517
; Publication No. US20050256299A1
; GENERAL INFORMATION:
; APPLICANT: University of Sheffield
; APPLICANT: Biosynex Incorporated
; APPLICANT: Foster, Simon
; APPLICANT: Mond, James
; TITLE OF INVENTION: Antigenic Polypeptides
; FILE REFERENCE: P100629W0
; CURRENT APPLICATION NUMBER: US/10/485,517
; CURRENT FILING DATE: 2004-02-02
; PRIOR APPLICATION NUMBER: GB 0118825.9
; PRIOR FILING DATE: 2001-08-02
; PRIOR APPLICATION NUMBER: GB 0200349.9
; PRIOR FILING DATE: 2002-01-09
; NUMBER OF SEQ ID NOS: 424
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 145
; LENGTH: 1107
; TYPE: PRT
; ORGANISM: Staphylococcus aureus
US-10-485-517-145

Query Match      39.6%; Score 38; DB 6; Length 1107;
Best Local Similarity 45.0%; Pred. No. 1.5e+02;
Matches 9; Conservative 4; Mismatches 3; Indels 4; Gaps 1;

Qy      5 INNL----TAQQQAAQAVN 20
Db      711 LNNLSNLTTPQQALENQIN 730

RESULT 13
US-10-793-626-3188
; Sequence 3188, Application US/10793626
; Publication No. US20050255478A1
; GENERAL INFORMATION:
; APPLICANT: KIMMERLY, WILLIAM JOHN
; TITLE OF INVENTION: STAPHYLOCOCCUS EPIDERMIDIS NUCLEIC ACIDS AND PROTEINS
; FILE REFERENCE: PU3480US
; CURRENT APPLICATION NUMBER: US/10/793,626
; CURRENT FILING DATE: 2004-03-04
; PRIOR APPLICATION NUMBER: 60/164,258
; PRIOR FILING DATE: 1999-11-09
; NUMBER OF SEQ ID NOS: 4472
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 3188
; LENGTH: 1279
; TYPE: PRT
; ORGANISM: Artificial Sequence

; CURRENT APPLICATION NUMBER: US/10/467,657
; CURRENT FILING DATE: 2003-08-11
; PRIOR APPLICATION NUMBER: GB-0103424.8
; PRIOR FILING DATE: 2001-02-12
; NUMBER OF SEQ ID NOS: 9218
; SOFTWARE: SeqWin99, version 1.04
; SEQ ID NO 3988
; LENGTH: 612
; TYPE: PRT
; ORGANISM: Neisseria gonorrhoeae
US-10-467-657-3988

Query Match      40.6%; Score 39; DB 6; Length 612;
Best Local Similarity 47.4%; Pred. No. 51;
Matches 9; Conservative 3; Mismatches 7; Indels 0; Gaps 0;

Qy      1 QDSKNNLTAQQQAAQAV 19
Db      351 QEQKNNELARLKKIQNV 369

RESULT 12
US-10-485-517-145
; Sequence 145, Application US/10485517
; Publication No. US20050256299A1
; GENERAL INFORMATION:
; APPLICANT: University of Sheffield
; APPLICANT: Biosynex Incorporated
; APPLICANT: Foster, Simon
; APPLICANT: Mond, James
; TITLE OF INVENTION: Antigenic Polypeptides
; FILE REFERENCE: P100629W0
; CURRENT APPLICATION NUMBER: US/10/485,517
; CURRENT FILING DATE: 2004-02-02
; PRIOR APPLICATION NUMBER: GB 0118825.9
; PRIOR FILING DATE: 2001-08-02
; PRIOR APPLICATION NUMBER: GB 0200349.9
; PRIOR FILING DATE: 2002-01-09
; NUMBER OF SEQ ID NOS: 424
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 145
; LENGTH: 1107
; TYPE: PRT
; ORGANISM: Staphylococcus aureus
US-10-485-517-145

Query Match      39.6%; Score 38; DB 6; Length 1107;
Best Local Similarity 45.0%; Pred. No. 1.5e+02;
Matches 9; Conservative 4; Mismatches 3; Indels 4; Gaps 1;

Qy      5 INNL----TAQQQAAQAVN 20
Db      711 LNNLSNLTTPQQALENQIN 730

RESULT 13
US-10-793-626-3188
; Sequence 3188, Application US/10793626
; Publication No. US20050255478A1
; GENERAL INFORMATION:
; APPLICANT: KIMMERLY, WILLIAM JOHN
; TITLE OF INVENTION: STAPHYLOCOCCUS EPIDERMIDIS NUCLEIC ACIDS AND PROTEINS
; FILE REFERENCE: PU3480US
; CURRENT APPLICATION NUMBER: US/10/793,626
; CURRENT FILING DATE: 2004-03-04
; PRIOR APPLICATION NUMBER: 60/164,258
; PRIOR FILING DATE: 1999-11-09
; NUMBER OF SEQ ID NOS: 4472
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 3188
; LENGTH: 1279
; TYPE: PRT
; ORGANISM: Artificial Sequence

; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: synthetic
; OTHER INFORMATION: amino acid sequence
US-10-793-626-3188

Query Match      38.5%; Score 37; DB 6; Length 1279;
Best Local Similarity 43.8%; Pred. No. 2.5e+02;
Matches 7; Conservative 4; Mismatches 5; Indels 0; Gaps 0;

Qy      5 INNLTAQQQAAQAVN 20
Db      35 INRVSAELARVQAQIN 50

RESULT 14
US-11-069-834-48
; Sequence 48, Application US/11069834
; Publication No. US20050276811A1
; GENERAL INFORMATION:
; APPLICANT: CARROLL, MICHAEL C.
; APPLICANT: MOORE JR., FRANCIS D.
; APPLICANT: HECHTMAN, HERBERT B.
; TITLE OF INVENTION: NATURAL IGM ANTIBODIES AND INHIBITORS THEREOF
; FILE REFERENCE: CRA-002.01
; CURRENT APPLICATION NUMBER: US/11/069,834
; CURRENT FILING DATE: 2005-03-01
; PRIOR APPLICATION NUMBER: 60/588,648
; PRIOR FILING DATE: 2004-07-16
; PRIOR APPLICATION NUMBER: 60/549,123
; PRIOR FILING DATE: 2004-03-01
; NUMBER OF SEQ ID NOS: 65
; SOFTWARE: PatentIn Ver. 3.3
; SEQ ID NO 48
; LENGTH: 1960
; TYPE: PRT
; ORGANISM: Mus musculus
US-11-069-834-48

Query Match      38.5%; Score 37; DB 7; Length 1960;
Best Local Similarity 38.9%; Pred. No. 4.1e+02;
Matches 7; Conservative 6; Mismatches 5; Indels 0; Gaps 0;

Qy      2 DSKNNLTAQQQAAQAV 19
Db      1820 BEQLDNETKERQAASKQV 1837

RESULT 15
US-10-793-626-3154
; Sequence 3154, Application US/10793626
; Publication No. US20050255478A1
; GENERAL INFORMATION:
; APPLICANT: KIMMERLY, WILLIAM JOHN
; TITLE OF INVENTION: STAPHYLOCOCCUS EPIDERMIDIS NUCLEIC ACIDS AND PROTEINS
; FILE REFERENCE: PU3480US
; CURRENT APPLICATION NUMBER: US/10/793,626
; CURRENT FILING DATE: 2004-03-04
; PRIOR APPLICATION NUMBER: 60/164,258
; PRIOR FILING DATE: 1999-11-09
; NUMBER OF SEQ ID NOS: 4472
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 3154
; LENGTH: 1095
; TYPE: PRT
; ORGANISM: Artificial Sequence

; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: synthetic
; OTHER INFORMATION: amino acid sequence
US-10-793-626-3154

Query Match      38.0%; Score 36.5; DB 6; Length 1095;
Best Local Similarity 47.4%; Pred. No. 2.5e+02;
Matches 9; Conservative 4; Mismatches 5; Indels 1; Gaps 1;
```

Qy 1 QDSKNNLTAQQAQV 19
DB 513 QDS-INNYVAKRNEQSAI 530

Search completed: January 25, 2006, 20:11:08
Job time : 5.5 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.1.6
Copyright (c) 1993 - 2006 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: January 25, 2006, 18:48:24 ; Search time 68.2 Seconds
(without alignments)
128.850 Million cell updates/sec

Title: US-10-797-821-8
Perfect score: 93
Sequence: 1 QQAQAQVNTICQVSLQT 20

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 2443163 seqs, 439378781 residues

Total number of hits satisfying chosen parameters: 2443163

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : A_Geneseq_21.*
1: geneseqp1980s.*
2: geneseqp1990s.*
3: geneseqp2000s.*
4: geneseqp2001s.*
5: geneseqp2002s.*
6: geneseqp2003as.*
7: geneseqp2003bs.*
8: geneseqp2004s.*
9: geneseqp2005s.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	93	100.0	20	7	ADD93628 Streptococcus
2	93	100.0	20	9	ADK37251 Streptococcus
3	93	100.0	431	7	ADD93649 Streptococcus
4	93	100.0	431	7	ADD93650 Streptococcus
5	93	100.0	431	7	ADD93653 Streptococcus
6	93	100.0	431	9	ADK37272 Streptococcus
7	93	100.0	431	9	ADK37273 Streptococcus
8	93	100.0	431	9	ADK37276 Streptococcus
9	93	100.0	431	9	AE91500 Microbial
10	93	100.0	432	7	ADD93651 Streptococcus
11	93	100.0	432	7	ADD93652 Streptococcus
12	93	100.0	432	9	ADK37274 Streptococcus
13	93	100.0	432	9	ADK37275 Streptococcus
14	75	80.6	20	7	ADD93629 Streptococcus
15	75	80.6	20	9	ADK37252 Streptococcus
16	58	62.4	398	5	ABP25919 Streptococcus
17	58	62.4	398	8	ADR83884 Streptococcus
18	58	62.4	398	8	AE91542 Microbial
19	57	61.3	210	2	AAV22579 Bacterial
20	57	61.3	392	6	ABU02747 S. pneumoniae
21	57	61.3	392	8	ADK47859 Streptococcus
22	57	61.3	392	8	ADT50227 Streptococcus
23	57	61.3	392	8	ADT50226 Streptococcus
24	57	61.3	392	8	ADT50165 Streptococcus

25	57	61.3	399	8	ADR94595	Adt94595 Novel S.
26	57	61.3	399	9	AEA58465	Aea58465 Streptococcus
27	55	59.1	447	5	ABP29684	Abp29684 Streptococcus
28	55	59.1	447	8	ADU69524	Adu69524 S agalactiae
29	55	59.1	447	8	ADV88392	Adv88392 Streptococcus
30	55	59.1	447	8	ADV81808	Adv81808 Streptococcus
31	55	59.1	447	8	ADV79645	Adv79645 Streptococcus
32	51	54.8	20	7	ADD93622	Add93622 Streptococcus
33	51	54.8	20	9	ADK37245	Adk37245 Streptococcus
34	50	53.8	1024	5	ABB04822	Abb04822 LDL recep
35	48	51.6	603	4	ACG71388	Aag71388 Human gen
36	48	51.6	825	8	ADS44071	Ads44071 Bacterial
37	48	51.6	1024	5	ABB04818	Abb04818 LDL recep
38	48	51.6	1024	5	ABB04824	Abb04824 LDL recep
39	48	51.6	2540	9	ADW11843	Adw11843 Talin ful
40	48	51.6	2541	3	AAB41087	Aab41087 Human ORF
41	48	51.6	2541	5	ABB81459	Abb81459 Human Tal
42	48	51.6	2541	5	ADG90450	Adg90450 Human tal
43	48	51.6	2541	6	ABR47614	AbR47614 Breast ca
44	48	51.6	2541	7	ADE58866	Ade58866 Rat Proce
45	48	51.6	2541	7	ADD45522	Add45522 Human Pro

ALIGNMENTS

RESULT 1
ADD93628
ID ADD93628 standard; peptide; 20 AA.
XX AC ADD93628;
XX AC ADD93628;
DT 29-JAN-2004 (first entry)
XX DE Streptococcus mutans glucan binding protein-B peptide fragment.
XX KW Glucan binding protein-B; GbpB; vaccine; epitope; immunogen.
XX OS Streptococcus mutans.
XX PN WO2003075845-A2.
XX PD 18-SEP-2003.
XX PF 07-MAR-2003; 2003WO-US006962.
XX PR 07-MAR-2002; 2002US-0363209P.
XX PR 08-AUG-2002; 2002US-0402483P.
XX (FORS-) FORSYTH INST.
XX Smith DJ, Taubman MA;
XX WPI; 2003-845091/78.
Composition useful as vaccines for dental caries comprises a fragment of a glucan binding protein-B binding to a major histocompatibility complex class II protein.

Claim 11; Page 10; 49pp; English.

The present sequence is that of a peptide comprising amino acid residues 48-67 of the glucan binding protein-B (GbpB) of Streptococcus mutans strain SK32 ADD93649. The peptide binds to a major histocompatibility complex (MHC) class II protein. It was identified as a potential B cell epitope using a matrix-based algorithm for epitope prediction, which was used to search the primary amino acid sequence of GbpB for known MHC class II binding motifs. The peptide can be used in immunogenic compositions and subunit vaccines for dental caries. These comprise MHC class II protein-binding GbpB peptides covalently linked with peptide subunits (preferably from the catalytic domain) of a glucosyltransferase. The compositions are used in a claimed method of eliciting production of an antibody in a mammal. Dieptopic or multieptopic polypeptides can be

XX Composition useful as vaccines for dental caries comprises a fragment of
 PT a glucan binding protein-B binding to a major histocompatibility complex
 PT class II protein.

XX Claim 5; Page 8; 49pp; English.

XX The present sequence is the protein sequence of the glucan binding
 CC protein-B (GbpB) of Streptococcus mutans strain 15P2. The sequence
 CC includes a number of human leucocyte antigen (HLA)-binding epitopes. The
 CC invention provides immunogenic compositions and vaccines for dental
 CC caries. The compositions comprise major histocompatibility complex (MHC)
 CC class II protein-binding GbpB peptides, especially HLA-binding peptides,
 CC covalently linked with peptide subunits of a glucosyltransferase. The
 CC compositions are used in a claimed method of eliciting production of an
 CC antibody in a mammal. Dieptopic or multiepitopic polypeptides can be
 CC prepared synthetically or by recombinant DNA technology. Antibodies
 CC raised against MHC class II binding fragments of GbpB can be used in
 CC passive immunisation.

XX Sequence 432 AA;

Query Match 100.0%; Score 93; DB 7; Length 432;
 Best Local Similarity 100.0%; Pred. No. 1.8e-06;
 Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 QQAQAQVNTTIGQVVSALQT 20

Db 48 QQAQAQVNTTIGQVVSALQT 67

RESULT 11

ADD93652
 ID ADD93652 standard; protein; 432 AA.

XX ADD93652;

DT 29-JAN-2004 (first entry)

DE Streptococcus mutans glucan binding protein-B.

XX Glucan binding protein-B; GbpB; vaccine; anticaries; epitope; immunogen.

XX Streptococcus mutans.

XX WO2003075845-A2.

XX 18-SEP-2003.

XX 07-MAR-2003; 2003WO-US006962.

XX 07-MAR-2002; 2002US-0363209P.

XX 08-AUG-2002; 2002US-0402483P.

XX (FORS-) FORSYTH INST.

XX Smith DJ, Taubman MA;

XX WPI; 2003-845091/78.

XX GENBANK; AY046413.

XX Composition useful as vaccines for dental caries comprises a fragment of
 PT a glucan binding protein-B binding to a major histocompatibility complex
 PT class II protein.

XX Claim 5; Page 8; 49pp; English.

XX The present sequence is the protein sequence of the glucan binding
 CC protein-B (GbpB) of Streptococcus mutans strain 3SN1. The sequence
 CC includes a number of human leucocyte antigen (HLA)-binding epitopes. The
 CC invention provides immunogenic compositions and vaccines for dental
 CC caries. The compositions comprise major histocompatibility complex (MHC)
 CC class II protein-binding GbpB peptides, especially HLA-binding peptides,

CC covalently linked with peptide subunits of a glucosyltransferase. The
 CC compositions are used in a claimed method of eliciting production of an
 CC antibody in a mammal. Dieptopic or multiepitopic polypeptides can be
 CC prepared synthetically or by recombinant DNA technology. Antibodies
 CC raised against MHC class II binding fragments of GbpB can be used in
 CC passive immunisation.

XX Sequence 432 AA;

Query Match 100.0%; Score 93; DB 7; Length 432;
 Best Local Similarity 100.0%; Pred. No. 1.8e-06;
 Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 QQAQAQVNTTIGQVVSALQT 20

Db 48 QQAQAQVNTTIGQVVSALQT 67

RESULT 12

ADX37274
 ID ADX37274 standard; protein; 432 AA.

XX ADX37274;

DT 21-APR-2005 (first entry)

DE Streptococcus mutans glucan binding protein B variant #3.

XX immunogenicity; immune stimulation; glucan binding protein-B;
 KW microparticle; major histocompatibility complex; tooth disease.

XX Streptococcus mutans.

XX US2005031633-A1.

XX 10-FEB-2005.

XX 09-MAR-2004; 2004US-00797821.

XX 13-APR-1998; 98US-0081550P.

XX 08-JAN-1999; 99US-0115142P.

XX 12-APR-1999; 99US-00290049.

XX 07-MAR-2002; 2002US-0363209P.

XX 08-AUG-2002; 2002US-0402483P.

XX 07-MAR-2003; 2003US-00383930.

XX (SMIT/) SMITH D J.

XX (TAUB/) TAUBMAN M A.

XX Smith DJ, Taubman MA;

XX WPI; 2005-151644/16.

XX New composition comprising a fragment of a glucan binding protein-B
 PT (GbpB) that binds to MHC class II protein, and a biocompatible
 PT microparticle, useful for producing an antibody (claimed) for immunizing
 PT mammals against dental caries.

XX Claim 3; SEQ ID NO 31; 73pp; English.

XX The invention relates to a composition comprising a fragment of a glucan
 CC binding protein-B (GbpB) and a biocompatible microparticle, where the
 CC fragment binds to a major histocompatibility complex (MHC) class II
 CC protein. The composition is useful for producing an antibody for
 CC immunizing mammals against dental caries. This sequence corresponds to a
 CC Streptococcus mutans GbpB protein of the invention.

XX Sequence 432 AA;

Query Match 100.0%; Score 93; DB 9; Length 432;
 Best Local Similarity 100.0%; Pred. No. 1.8e-06;
 Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

QY 1 QQAQAQVNTIQGVSAALQT 20
Db 48 QQAQAQVNTIQGVSAALQT 67

RESULT 13
ADX37275
ID ADX37275 standard; protein; 432 AA.
XX
XX AC ADX37275;
XX
XX DT 21-APR-2005 (first entry)
XX
XX DE Streptococcus mutant glucan binding protein B variant #4.
XX
XX KW immunogenicity; immune stimulation; glucan binding protein-B;
KW microparticle; major histocompatibility complex; tooth disease.
XX
XX OS Streptococcus mutans.
XX
XX PN US2005031633-A1.
XX
XX PD 10-FEB-2005.
XX
XX PF 09-MAR-2004; 2004US-00797821.
XX
XX PR 13-APR-1998; 98US-0081550P.
PR 08-JAN-1999; 99US-0115142P.
PR 12-APR-1999; 99US-00290049.
PR 07-MAR-2002; 2002US-0363209P.
PR 08-AUG-2002; 2002US-0402483P.
PR 07-MAR-2003; 2003US-00383930.
XX
XX PA (SMIT/) SMITH D J.
XX PA (TAUB/) TAUBMAN M A.
XX
XX PI Smith DJ, Taubman MA;
XX
XX WPI; 2005-151644/16.
XX
XX DR New composition comprising a fragment of a glucan binding protein-B
XX (GbpB) that binds to MHC class II protein, and a biocompatible
XX microparticle, useful for producing an antibody (claimed) for immunizing
XX mammals against dental caries.
XX
XX PS Claim 3; SEQ ID NO 32; 73pp; English.
XX
XX CC The invention relates to a composition comprising a fragment of a glucan
XX binding protein-B (GbpB) and a biocompatible microparticle, where the
XX fragment binds to a major histocompatibility complex (MHC) class II
XX protein. The composition is useful for producing an antibody for
XX immunizing mammals against dental caries. This sequence corresponds to a
XX Streptococcus mutans GbpB protein of the invention.
XX
XX SQ Sequence 432 AA;

Query Match 100.0%; Score 93; DB 9; Length 432;
Best Local Similarity 100.0%; Pred. No. 1.8e-06;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 QQAQAQVNTIQGVSAALQT 20
Db 48 QQAQAQVNTIQGVSAALQT 67

RESULT 14
ADD93629
ID ADD93629 standard; peptide; 20 AA.
XX
XX AC ADD93629;
XX
XX DT 29-JAN-2004 (first entry)
XX
XX

```

```

DE Streptococcus mutans glucan binding protein-B peptide fragment.
XX
KW Glucan binding protein-B; GbpB; vaccine; anticaries; epitope; immunogen.
XX
XX OS Streptococcus mutans.
XX
XX PN WO2003075845-A2.
XX
XX PD 18-SEP-2003.
XX
XX PF 07-MAR-2003; 2003WO-US006962.
XX
XX PR 07-MAR-2002; 2002US-0363209P.
PR 08-AUG-2002; 2002US-0402483P.
XX
XX PA (FORS-) FORSYTH INST.
XX
XX PI Smith DJ, Taubman MA;
XX
XX WPI; 2003-845091/78.
XX
XX CC Composition useful as vaccines for dental caries comprises a fragment of
XX a glucan binding protein-B binding to a major histocompatibility complex
XX class II protein.
XX
XX PS Claim 11; Page 10; 49pp; English.
XX
XX CC The present sequence is that of a peptide comprising amino acid residues
XX 52-71 of the glucan binding protein-B (GbpB) of Streptococcus mutans
XX strain SK32 ADD93649. The peptide binds to a major histocompatibility
XX complex (MHC) class II protein. It was identified as a potential B cell
XX epitope using a matrix-based algorithm for epitope prediction, which was
XX used to search the primary amino acid sequence of GbpB for known MHC
XX class II binding motifs. The peptide can be used in immunogenic
XX compositions and subunit vaccines for dental caries. These comprise MHC
XX class II protein-binding GbpB peptides covalently linked with peptide
XX subunits (preferably from the catalytic domain) of a glucosyltransferase.
XX The compositions are used in a claimed method of eliciting production of
XX an antibody in a mammal. Diepitopic or multipitopic polypeptides can be
XX prepared synthetically or by recombinant DNA technology. Antibodies
XX raised against MHC class II binding fragments of GbpB can be used in
XX passive immunisation.
XX
XX SQ Sequence 20 AA;

Query Match 80.6%; Score 75; DB 7; Length 20;
Best Local Similarity 100.0%; Pred. No. 4.5e-05;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 5 QAQVNTIQGVSAALQT 20
Db 1 QAQVNTIQGVSAALQT 16

RESULT 15
ADX37252
ID ADX37252 standard; peptide; 20 AA.
XX
XX AC ADX37252;
XX
XX DT 21-APR-2005 (first entry)
XX
XX DE Streptococcus mutant glucan binding protein B peptide #9.
XX
XX KW immunogenicity; immune stimulation; glucan binding protein-B;
KW microparticle; major histocompatibility complex; tooth disease.
XX
XX OS Streptococcus mutans.
XX
XX PN US2005031633-A1.
XX
XX PD 10-FEB-2005.
XX

```


GenCore version 5.1.6
Copyright (c) 1993 - 2006 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: January 25, 2006, 18:48:55 ; Search time 9.1 Seconds
(without alignments)
211.465 Million cell updates/sec

Title: US-10-797-821-8
Perfect score: 93
Sequence: 1 QQAQAQVNTIQGVSAIQ 20
Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283416 seqs, 96216763 residues
Total number of hits satisfying chosen parameters: 283416
Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : PIR 80: *
1: pir1: *
2: pir2: *
3: pir3: *
4: pir4: *

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	57	61.3	392	2 G95258	secreted 45 kd pro
2	57	61.3	392	2 B98124	general stress pro
3	57	61.3	507	2 S05542	hypothetical prote
4	48	51.6	825	2 S62042	probable membrane
5	48	51.6	2541	2 S11661	talin - mouse
6	46	49.5	259	2 G82865	conjugal transfer
7	46	49.5	454	2 AC0498	probable thiophene
8	46	49.5	456	2 E86903	hypothetical prote
9	46	49.5	461	2 JN0097	secreted 45K prote
10	46	49.5	1379	2 T13718	pollux gene protei
11	46	49.5	2427	2 T16613	hypothetical prote
12	45	48.4	445	2 B40590	flagellar distal c
13	45	48.4	562	2 S75308	DNA ligase (EC 6.5
14	45	48.4	734	2 AH2359	hypothetical prote
15	44	47.3	125	2 S74966	hypothetical prote
16	44	47.3	393	2 S64733	retrovirus-related
17	44	47.3	434	2 AH1993	hypothetical prote
18	44	47.3	472	2 S08509	cruciferin precurs
19	44	47.3	488	2 B35540	cruciferin 2/3 pre
20	43.5	46.8	793	1 S60735	splicing factor SF
21	43	46.2	88	2 A38085	S-layer glycoprote
22	43	46.2	234	2 T00931	hypothetical prote
23	43	46.2	454	2 C86055	hypothetical prote
24	43	46.2	454	2 AE0957	thiophene and fura
25	43	46.2	454	2 A98209	hypothetical prote
26	43	46.2	454	2 C65173	thiophene and fura
27	43	46.2	827	2 A37849	S-layer protein -
28	43	46.2	1474	2 T20488	hypothetical clone
29	42	45.2	112	2 S26221	cruciferin (clone

30	42	45.2	233	2 S26222	cruciferin (clone
31	42	45.2	390	2 A72108	hypothetical prote
32	42	45.2	390	2 B86514	hypothetical prote
33	42	45.2	439	2 AD2555	hypothetical prote
34	42	45.2	460	2 I57546	Rabin3 - rat
35	42	45.2	496	2 S25091	cruciferin Bnc2 -
36	42	45.2	528	2 B75310	conserved hypotet
37	42	45.2	902	2 G83635	probable C1pA/B-ty
38	41	44.1	90	2 A10868	conserved hypotet
39	41	44.1	90	2 G90992	hypothetical prote
40	41	44.1	90	2 B85838	hypothetical prote
41	41	44.1	90	2 H64977	hypothetical prote
42	41	44.1	142	2 C81359	hypothetical prote
43	41	44.1	154	2 G84396	hypothetical prote
44	41	44.1	233	2 S34956	gene 26 protein -
45	41	44.1	522	2 B42804	immunodominant typ

ALIGNMENTS

RESULT 1
G95258
secreted 45 kd protein [imported] - Streptococcus pneumoniae (strain TIGR4)
C:Species: Streptococcus pneumoniae
C:Date: 03-Aug-2001 #sequence_revision 03-Aug-2001 #text_change 09-Jul-2004
C:Accession: G95258
R:Tetzelin, H.; Nelson, K.E.; Paulsen, I.T.; Eisen, J.A.; Read, T.D.; Peterson, S.; Hei
on, J.D.; Unayam, L.A.; White, O.; Salzberg, S.L.; Lewis, M.R.; Radune, D.; Holtzapple,
nson, T.; Hickey, E.K.; Holt, I.E.
Science 293, 498-506, 2001
A:Authors: Loftus, B.J.; Yang, F.; Smith, H.O.; Venter, J.C.; Dougherty, B.A.; Morrison,
A:Title: Complete Genome Sequence of a virulent isolate of Streptococcus pneumoniae.
A:Reference number: A95000; MUID:21357209; PMID:11463916
A:Accession: G95258
A>Status: Preliminary
A:Molecule type: DNA
A:Residues: 1-392 <KUR>
A:Cross-references: UNIPROT:Q97N55; UNIPROT:Q8DMY4; UNIPARC:UPI0000051B81; GB:AE005672;
A:Experimental source: strain TIGR4
C:Genetics:
A:Gene: SP2216

Query Match 61.3%; Score 57; DB 2; Length 392;
Best Local Similarity 68.4%; Pred. No. 0.14;
Matches 13; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

QY 1 QQAQAQVNTIQGVSAIQ 19
DB 48 QQAQAQVNTIQGVSAIQ 66

RESULT 2
B98124
General stress protein GSP-781 [imported] - Streptococcus pneumoniae (strain R6)
C:Species: Streptococcus pneumoniae
C:Date: 22-Oct-2001 #sequence_revision 22-Oct-2001 #text_change 09-Jul-2004
C:Accession: B98124
R:Hoskins, J.A.; Alborn Jr., W.; Arnold, J.; Blaszcak, L.; Burgett, S.; DeHoff, B.S.; ;
e, R.; LeBlanc, D.J.; Lee, L.N.; Lefkowitz, E.J.; Lu, J.; Matsushima, P.; McAhren, S.; ;
Y, P.; Sun, P.M.; Winkler, M.E.
J. Bacteriol. 183, 5709-5717, 2001
A:Authors: Yang, Y.; Young-Bellido, M.; Zhao, G.; Zook, C.; Baltz, R.H.; Jaskunas, S.R.
A:Title: Genome of the Bacterium Streptococcus pneumoniae Strain R6.
A:Reference number: A97872; MUID:21429245; PMID:11544234
A:Accession: B98124
A>Status: Preliminary
A:Molecule type: DNA
A:Residues: 1-392 <KUR>
A:Cross-references: UNIPROT:Q97N55; UNIPROT:Q8DMY4; UNIPARC:UPI0000051B81; GB:AE007317;
C:Genetics:
A:Gene: gsp-781

A>Title: Genome sequence of *Yersinia pestis*, the causative agent of plague.

A;Reference number: AB0001; MUID:21470413; PMID:11586360

A;Accession: AC0498

A;Status: preliminary

A;Molecule type: DNA

A;Residues: 1-454 <KUR>

A;Cross-references: UNIPROT:Q8Z9U2; UNIPARC:UPI000013746C; GB:AL590842; PIDN:CAC93552.1;

C;Genetics:

A;Gene: chdF

C;Superfamily: thiophen / furan oxidation protein; translation elongation factor Tu homolog

Query Match 49.5%; Score 46; DB 2; Length 454;

Best Local Similarity 52.9%; Pred. No. 10;

Matches 9; Conservative 5; Mismatches 3; Indels 0; Gaps 0;

QY 1 QQAQAQVNTTGGQVSA 17

:|||::||::|||

Db 138 EQAARSVNSLQGAFA 154

RESULT 8

EB6903

hypothetical protein usp45 [imported] - *Lactococcus lactis* subsp. *lactis* (strain IL1403)

C;Species: *Lactococcus lactis* subsp. *lactis*

C;Date: 23-Mar-2001 #sequence_revision 23-Mar-2001 #text_change 09-Jul-2004

C;Accession: EB6903

R;Bolotin, A.; Wincker, P.; Mauger, S.; Jaillon, O.; Malarme, K.; Weissenbach, J.; Ehrlich

Genome Res. 11, 731-753, 2001

A>Title: The complete genome sequence of the lactic acid bacterium *Lactococcus lactis* se

A;Reference number: A86625; MUID:21235186; PMID:11337471

A;Accession: EB6903

A;Status: preliminary

A;Molecule type: DNA

A;Residues: 1-456 <STO>

A;Cross-references: UNIPROT:Q9CDJ1; UNIPARC:UPI000006BFC; GB:AE005176; PID:g12725296; E

C;Genetics:

A;Gene: usp45

Query Match 49.5%; Score 46; DB 2; Length 456;

Best Local Similarity 56.2%; Pred. No. 11;

Matches 9; Conservative 5; Mismatches 2; Indels 0; Gaps 0;

QY 4 AQAQVNTTGGQVSA 19

|||||::||::|||

Db 50 AQAQVDSLQSKVDSLQ 65

RESULT 9

JN0097

secreted 45K protein precursor - *Lactococcus lactis*

C;Species: *Lactococcus lactis*

C;Date: 16-Jul-1999 #sequence_revision 16-Jul-1999 #text_change 16-Jul-1999

C;Accession: JN0097

R;van Asseeldonk, M.; Rutten, G.; Otteman, M.; Sierzen, R.J.; de Vos, W.M.; Simons, G.

Gene 95, 155-160, 1990

A>Title: Cloning of usp45, a gene encoding a secreted protein from *Lactococcus lactis* su

A;Reference number: JN0097; MUID:91071599; PMID:2123812

A;Accession: JN0097

A;Status: preliminary

A;Molecule type: DNA

A;Residues: 1-461 <VAN>

A;Cross-references: UNIPARC:UPI000016D72B; GB:M35374

Query Match 49.5%; Score 46; DB 2; Length 461;

Best Local Similarity 56.2%; Pred. No. 11;

Matches 9; Conservative 5; Mismatches 2; Indels 0; Gaps 0;

QY 4 AQAQVNTTGGQVSA 19

|||||::||::|||

Db 50 AQAQVDSLQSKVDSLQ 65

RESULT 10

TI3718

pollux gene protein - fruit fly (*Drosophila melanogaster*)

C;Species: *Drosophila melanogaster*

C;Date: 13-Aug-1999 #sequence_revision 13-Aug-1999 #text_change 09-Jul-2004

C;Accession: TI3718

R;Xu, X.Z.; Wes, P.D.; Chen, H.; Li, H.S.; Yu, M.; Morgan, S.; Liu, Y.; Montell, C.

J. Biol. Chem. 273, 31297, 1998

A>Title: Retinal targets for calmodulin include proteins implicated in synaptic transmi:

A;Reference number: Z17709; MUID:99030403; PMID:9813038

A;Accession: TI3718

A;Status: preliminary; translated from GB/EMBL/DBJ

A;Molecule type: mRNA

A;Residues: 1-1379 <XUX>

A;Cross-references: UNIPROT:Q9TY11; UNIPARC:UPI0000084062; EMBL:Y17919; NID:g3893102; P

C;Genetics:

A;Gene: pollux

A;Cross-references: FlyBase:FBgn0004879

Query Match 49.5%; Score 46; DB 2; Length 1379;

Best Local Similarity 42.1%; Pred. No. 33;

Matches 8; Conservative 6; Mismatches 5; Indels 0; Gaps 0;

QY 1 QQAQAQVNTTGGQVSA 19

:|||::||::|||

Db 1051 RSSQAQITTLQSQVQSLQ 1069

RESULT 11

TI6613

hypothetical protein K10B3.10 - *Caenorhabditis elegans*

C;Species: *Caenorhabditis elegans*

C;Date: 20-Sep-1999 #sequence_revision 20-Sep-1999 #text_change 09-Jul-2004

C;Accession: TI6613

R;Gattung, S.

submitted to the EMBL Data Library, February 1996

A;Description: The sequence of *C. elegans* cosmid K10B3.

A;Reference number: Z18546

A;Accession: TI6613

A;Status: preliminary; translated from GB/EMBL/DBJ

A;Molecule type: DNA

A;Residues: 1-2427 <GAT>

A;Cross-references: UNIPROT:Q21408; UNIPARC:UPI000007E8A5; EMBL:U49941; NID:g1206038; P

A;Experimental source: strain Bristol N2; clone K10B3

C;Genetics:

A;Gene: CESP:K10B3.10

A;Map position: X

A;Introns: 68/2; 96/3; 138/3; 665/3; 1052/3; 1104/3; 1173/2; 1995/2; 2187/3; 2235/2; 233

C;Superfamily: spectrin alpha chain; calmodulin repeat homology; SH3 homology; spectrin/

C;Keywords: EF hand

F;990-1037/Domain: SH3 homology <SH3>

F;2050-2155/Domain: spectrin/dystrophin repeat homology <SPR>

F;2277-2309/Domain: calmodulin repeat homology <EFH>

Query Match 49.5%; Score 46; DB 2; Length 2427;

Best Local Similarity 52.6%; Pred. No. 60;

Matches 10; Conservative 4; Mismatches 5; Indels 0; Gaps 0;

QY 1 QQAQAQVNTTGGQVSA 19

|||||::||::|||

Db 1046 QQAQQQVNSIGGKQSEIE 1064

RESULT 12

B40590

flagellar distal capping protein homolog lafB - *Vibrio parahaemolyticus*

C;Species: *Vibrio parahaemolyticus*

C;Date: 12-May-1994 #sequence_revision 12-May-1994 #text_change 09-Jul-2004

C;Accession: B40590

R;McCarter, L.L.; Wright, M.E.

J. Bacteriol. 175, 3361-3371, 1993

A>Title: Identification of genes encoding components of the swarmer cell flagellar motor.

A;Reference number: A40590; MUID:93273702; PMID:8501040

GenCore version 5.1.6
Copyright (c) 1993 - 2006 Compugen Ltd.

OM protein - protein search, using sw model

Run on: January 25, 2006, 18:48:24 ; Search time 70.3 Seconds
(without alignments)
200.719 Million cell updates/sec

Title: US-10-797-821-8

Perfect score: 93

Sequence: 1 QQAQAQVNTTGGQVSALQT 20

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 2166443 seqs, 705528306 residues

Total number of hits satisfying chosen parameters: 2166443

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : UniProt_05.80.*

1: uniprot_sprot.*

2: uniprot_trembl.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	93	100.0	431	Q938V0_STRMU	Q938V0 streptococc
2	93	100.0	431	Q938V3_STRMU	Q938V3 streptococc
3	93	100.0	431	Q9AG98_STRMU	Q9AG98 streptococc
4	93	100.0	431	Q8DWM3_STRMU	Q8DWM3 streptococc
5	93	100.0	432	Q938V1_STRMU	Q938V1 streptococc
6	93	100.0	432	Q938V2_STRMU	Q938V2 streptococc
7	68	73.1	474	Q5M6K4_STRT2	Q5M6K4 streptococc
8	68	73.1	485	Q5M212_STRT1	Q5M212 streptococc
9	58	62.4	398	Q5XEL1_STRP6	Q5XEL1 streptococc
10	58	62.4	398	Q9A1Z8_STRPY	Q9A1Z8 streptococc
11	58	62.4	398	Q7CNQ7_STRP8	Q7CNQ7 streptococc
12	58	62.4	398	Q8P318_STRP3	Q8P318 streptococc
13	57	61.3	392	Q8DMY4_STRR6	Q8DMY4 streptococc
14	57	61.3	392	Q97N55_STRPN	Q97N55 streptococc
15	57	61.3	516	P54_ENTFC	P13692 enterococcu
16	57	61.3	576	Q9KJ3_ENTHR	Q9KJ3 enterococcu
17	56	60.2	211	Q9ZAS7_STRMU	Q9ZAS7 streptococc
18	55	59.1	118	Q8D167_ERWCT	Q8D167 erwinia car
19	55	59.1	447	Q9AKA4_STRAG	Q9AKA4 streptococc
20	55	59.1	447	Q8E2H1_STRAS	Q8E2H1 streptococc
21	55	59.1	447	Q8E7X9_STRAS	Q8E7X9 streptococc
22	50	53.8	524	Q63N34_BURPS	Q63N34 burkholderi
23	48	51.6	110	Q83BX1_COXBU	Q83BX1 coxiella bu
24	48	51.6	825	Q03088_YEAST	Q03088 saccharomyc
25	48	51.6	1209	Q84EC7_SYNPF	Q84EC7 synecococc
26	48	51.6	2429	Q8TCU6_HUMAN	Q8TCU6 homo sapien
27	48	51.6	2541	1_TLN1_HUMAN	Q9Y490 homo sapien
28	48	51.6	2541	1_TLN1_MOUSE	P26039 mus musculu
29	48	51.6	2541	2_Q86YD0_HUMAN	Q86YD0 homo sapien
30	48	51.6	2550	2_Q9UPX3_HUMAN	Q9UPX3 homo sapien
31	48	51.6	2564	2_Q80TM2_MOUSE	Q80TM2 mus musculu

32	47	50.5	180	2	Q4LJB1_9BURK	Q4LJB1 burkholderi
33	47	50.5	259	2	Q98LG5_RHILLO	Q98LG5 rhizobium 1
34	47	50.5	378	2	Q9N1X4_PIG	Q9N1X4 sus scrofa
35	47	50.5	579	2	Q4WYV3_ASPPU	Q4WYV3 aspergillus
36	47	50.5	759	2	Q53NW8_ORYZA	Q53NW8 oryza sativ
37	47	50.5	1172	2	Q7S348_NEUCR	Q7S348 neurospora
38	46	49.5	189	2	Q6SF50_9BACT	Q6SF50 uncultured
39	46	48.5	259	2	Q9PHG8_XYLFA	Q9PHG8 xylella fas
40	46	49.5	259	2	Q6MD00_PARUM	Q6MD00 parachlamyd
41	46	49.5	313	2	Q8PPA1_XANAC	Q8PPA1 xanthomonas
42	46	49.5	314	2	Q4UQX8_XANCP	Q4UQX8 xanthomonas
43	46	49.5	314	2	Q8PCJ3_XANCP	Q8PCJ3 xanthomonas
44	46	49.5	323	2	Q6N9H2_RHOPA	Q6N9H2 rhodopseudo
45	46	49.5	325	2	Q5GW48_XANOR	Q5GW48 xanthomonas

ALIGNMENTS

RESULT 1

Q938V0_STRMU Q938V0_STRMU PRELIMINARY; PRT; 431 AA.
 ID Q938V0;
 AC Q938V0;
 DT 01-DEC-2001 (Tremblrel. 19, Created)
 DT 01-DEC-2001 (Tremblrel. 19, Last sequence update)
 DT 01-MAR-2004 (Tremblrel. 26, Last annotation update)
 DE Glucan-binding protein B.
 OS Streptococcus mutans.
 OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
 OX Streptococcus.
 NCBI_TaxID=1309;
 RN [1]
 RP NUCLEOTIDE SEQUENCE.
 RC STRAIN=SSM3;
 RA Jin S., Duncan M.J., Taubman M.A., Smith D.J.;
 RT "Cloning of the gbpB gene from Streptococcus mutans.";
 RL J. Dent. Res. 79:224-224(2000).
 RN [2]
 RP NUCLEOTIDE SEQUENCE.
 RC STRAIN=SSM3;
 RX MEDLINE=21481971; PubMed=11598068;
 RX DOI=10.1128/JAI.69.11.6931-6941.2001;
 RA Mattos-Graner R.O., Jin S., King W.F., Chen T., Smith D.J.,
 Duncan M.J.;
 RT "Cloning of the Streptococcus mutans gene encoding glucan binding
 protein B and analysis of genetic diversity and protein production in
 clinical isolates.";
 RL Infect. Immun. 69:6931-6941(2001).
 DR EMBL; AY046414; AAK94504.1; -; Genomic_DNA.
 DR InterPro; IPR007921; CHAP.
 DR InterPro; IPR009148; Siba.
 DR Pfam; PF05257; CHAP; 1.
 DR PRINTS; PR01852; SIBAPROTEIN.
 DR PROSITE; PS50911; CHAP; 1.
 SQ SEQUENCE 431 AA; 44650 MW; 05D38D8B8C4609F CRC64;

Query Match 100.0%; Score 93; DB 2; Length 431;
 Best Local Similarity 100.0%; Pred. No. 2.3e-06;
 Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 QQAQAQVNTTGGQVSALQT 20
 |||||
 DB 48 QQAQAQVNTTGGQVSALQT 67

RESULT 2

Q938V3_STRMU Q938V3_STRMU PRELIMINARY; PRT; 431 AA.
 ID Q938V3;
 AC Q938V3;
 DT 01-DEC-2001 (Tremblrel. 19, Created)
 DT 01-DEC-2001 (Tremblrel. 19, Last sequence update)
 DT 01-MAR-2004 (Tremblrel. 26, Last annotation update)
 DE Glucan-binding protein B.

```

OS Streptococcus mutans.
OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
OC Streptococcus.
OX NCBI_TaxID=1309;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=5J32;
RA Jin S., Duncan M.J., Taubman M.A., Smith D.J.;
RT "Cloning of the gbpB gene from Streptococcus mutans.";
RL J. Dent. Res. 79:224-224 (2000).
RN [2]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=5J32;
RX MEDLINE=21481971; PubMed=11598068;
RX DOI=10.1128/IAI.69.11.6931-6941.2001;
RA Mattos-Graner R.O., Jin S., King W.F., Chen T., Smith D.J.,
RA Duncan M.J.;
RT "Cloning of the Streptococcus mutans gene encoding glucan binding
RT protein B and analysis of genetic diversity and protein production in
RT clinical isolates.";
RL Infect. Immun. 69:6931-6941 (2001).
DR EMBL; AF338445; AAK08104.1; -; Genomic_DNA.
DR EMBL; AY046410; AAK94500.1; -; Genomic_DNA.
DR InterPro; IPR007921; CHAP.
DR InterPro; IPR009148; Siba.
DR Pfam; PF05257; CHAP; 1.
DR PRINTS; PR01852; SIBAPROTEIN.
DR PROSITE; PS50911; CHAP; 1.
SQ SEQUENCE 431 AA; 44620 MW; 464FE3B563FB7E51 CRC64;

Query Match 100.0%; Score 93; DB 2; Length 431;
Best Local Similarity 100.0%; Pred. No. 2.3e-06;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 QQAQAQVNTTIGQVSALQT 20
DB 48 QQAQAQVNTTIGQVSALQT 67

RESULT 3
Q9AG98_STRMU
ID Q9AG98_STRMU PRELIMINARY; PRT; 431 AA.
AC Q9AG98_
DT 01-JUN-2001 (TrEMBLrel. 17, Created)
DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
DT 01-FEB-2005 (TrEMBLrel. 29, Last annotation update)
DE Immunodominant glycoprotein IDG-60 (Glucan-binding protein B).
GN Name=sagA;
OS Streptococcus mutans.
OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
OC Streptococcus.
OX NCBI_TaxID=1309;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=GS-5;
RX MEDLINE=21153617; PubMed=11254612;
RX DOI=10.1128/IAI.69.4.2493-2501.2001;
RA Chia J.S., Lee Y.Y., Huang P.T., Chen J.Y.;
RT "Identification of stress-responsive genes in Streptococcus mutans by
RT differential display reverse transcription-PCR.";
RL Infect. Immun. 69:2493-2501 (2001).
RN [2]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=GS-5;
RX MEDLINE=21481977; PubMed=11598074;
RX DOI=10.1128/IAI.69.11.6987-6998.2001;
RA Chia J.S., Chang L.Y., Shun C.T., Chang Y.Y., Chen J.Y.;
RT "A 60-kilodalton immunodominant glycoprotein is essential for cell
RT wall integrity and the maintenance of cell shape in Streptococcus
RT mutans.";
RL Infect. Immun. 69:6987-6998 (2001).
RN [3]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=GS-5;

```

```

RA Chia J.-S., Chang L.-Y., Lee Y.-Y., Chen J.-Y.;
RL Submitted (NOV-2001) to the EMBL/GenBank/DBJ databases.
RN [4]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=3VF4;
RA Jin S., Duncan M.J., Taubman M.A., Smith D.J.;
RT "Cloning of the gbpB gene from Streptococcus mutans.";
RL J. Dent. Res. 79:224-224 (2000).
RN [5]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=3VF4;
RX MEDLINE=21481971; PubMed=11598068;
RX DOI=10.1128/IAI.69.11.6931-6941.2001;
RA Mattos-Graner R.O., Jin S., King W.F., Chen T., Smith D.J.,
RA Duncan M.J.;
RT "Cloning of the Streptococcus mutans gene encoding glucan binding
RT protein B and analysis of genetic diversity and protein production in
RT clinical isolates.";
RL Infect. Immun. 69:6931-6941 (2001).
DR EMBL; AF338445; AAK08104.1; -; Genomic_DNA.
DR EMBL; AY046411; AAK94501.1; -; Genomic_DNA.
DR InterPro; IPR007921; CHAP.
DR InterPro; IPR009148; Siba.
DR Pfam; PF05257; CHAP; 1.
DR PRINTS; PR01852; SIBAPROTEIN.
DR PROSITE; PS50911; CHAP; 1.
SQ SEQUENCE 431 AA; 44592 MW; 3EBE21FC5E47232E CRC64;

Query Match 100.0%; Score 93; DB 2; Length 431;
Best Local Similarity 100.0%; Pred. No. 2.3e-06;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 QQAQAQVNTTIGQVSALQT 20
DB 48 QQAQAQVNTTIGQVSALQT 67

RESULT 4
Q8DWM3_STRMU
ID Q8DWM3_STRMU PRELIMINARY; PRT; 431 AA.
AC Q8DWM3_
DT 01-MAR-2003 (TrEMBLrel. 23, Created)
DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Putative secreted antigen GbpB/SagA; putative peptidoglycan
DE hydrolase.
GN Name=gbpB; OrderedLocusNames=SMU.22;
OS Streptococcus mutans.
OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
OC Streptococcus.
OX NCBI_TaxID=1309;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=UA159 / ATCC 700610 / Serotype C;
RX MEDLINE=23295063; PubMed=12397186; DOI=10.1073/pnas.172501299;
RX Ajdic D.J., McShan W.M., McLaughlin R.E., Savic G., Chang J.,
RA Carlson M.B., Primeaux C., Tian R., Kenton S., Jia H.G., Lin S.P.,
RA Qian Y., Li S., Zhu H., Najjar F.Z., Lai H., White J., Roe B.A.,
RA Ferretti J.J.;
RT "Genome sequence of Streptococcus mutans UA159, a cariogenic dental
RT pathogen.";
RL Proc. Natl. Acad. Sci. U.S.A. 99:14434-14439 (2002).
DR EMBL; AE014855; AAN57811.1; -; Genomic_DNA.
DR InterPro; IPR007921; CHAP.
DR InterPro; IPR009148; Siba.
DR Pfam; PF05257; CHAP; 1.
DR PRINTS; PR01852; SIBAPROTEIN.
DR PROSITE; PS50911; CHAP; 1.
KW Complete proteome.
SQ SEQUENCE 431 AA; 44620 MW; 2D1CA685248CCD3E CRC64;

Query Match 100.0%; Score 93; DB 2; Length 431;
Best Local Similarity 100.0%; Pred. No. 2.3e-06;

```

```

Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 QQAQAQVNTIQGVSAQT 20
DB 48 QQAQAQVNTIQGVSAQT 67

RESULT 5
Q338V1_STRMU
ID Q338V1_STRMU PRELIMINARY; PRT; 432 AA.
AC Q338V1;
DT 01-DEC-2001 (TrEMBLrel. 19, Created)
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Glucan-binding protein B.
OS Streptococcus mutans.
OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
OC Streptococcus.
OX NCBI_TaxID=1309;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=3SN1;
RA Jin S., Duncan M.J., Taubman M.A., Smith D.J.;
RT "Cloning of the gbpB gene from Streptococcus mutans.";
RL J. Dent. Res. 79:224-224(2000).
RN [2]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=3SN1;
RX MEDLINE=21481971; PubMed=11598068;
RX DOI=10.1128/JAI.69.11.6931-6941.2001;
RA Mattos-Graner R.O., Jin S., King W.F., Chen T., Smith D.J.,
RA Duncan M.J.;
RT "Cloning of the Streptococcus mutans gene encoding glucan binding
RT protein B and analysis of genetic diversity and protein production in
RT clinical isolates.";
RL Infect. Immun. 69:6931-6941(2001).
DR EMBL; AY046413; AAK94503.1; -; Genomic_DNA.
DR InterPro; IPR007921; CHAP.
DR Pfam; PF05257; CHAP; 1.
DR PRINTS; PR01852; SIBAPROTEIN.
DR PROSITE; PS50911; CHAP; 1.
SQ SEQUENCE 432 AA; 44652 MW; 3F88ECB9A1F3BB4F CRC64;

Query Match 100.0%; Score 93; DB 2; Length 432;
Best Local Similarity 100.0%; Pred. No. 2.3e-06;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 QQAQAQVNTIQGVSAQT 20
DB 48 QQAQAQVNTIQGVSAQT 67

RESULT 6
Q338V2_STRMU
ID Q338V2_STRMU PRELIMINARY; PRT; 432 AA.
AC Q338V2;
DT 01-DEC-2001 (TrEMBLrel. 19, Created)
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Glucan-binding protein B.
OS Streptococcus mutans.
OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
OC Streptococcus.
OX NCBI_TaxID=1309;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=15JP2;
RA Jin S., Duncan M.J., Taubman M.A., Smith D.J.;
RT "Cloning of the gbpB gene from Streptococcus mutans.";
RL J. Dent. Res. 79:224-224(2000).
RN [2]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=15JP2;
RX MEDLINE=21481971; PubMed=11598068;
RX DOI=10.1128/JAI.69.11.6931-6941.2001;
RA Mattos-Graner R.O., Jin S., King W.F., Chen T., Smith D.J.,
RA Duncan M.J.;
RT "Cloning of the Streptococcus mutans gene encoding glucan binding
RT protein B and analysis of genetic diversity and protein production in
RT clinical isolates.";
RL Infect. Immun. 69:6931-6941(2001).
DR EMBL; AY046413; AAK94503.1; -; Genomic_DNA.
DR InterPro; IPR007921; CHAP.
DR Pfam; PF05257; CHAP; 1.
DR PRINTS; PR01852; SIBAPROTEIN.
DR PROSITE; PS50911; CHAP; 1.
SQ SEQUENCE 432 AA; 44652 MW; 3F88ECB9A1F3BB4F CRC64;

Query Match 100.0%; Score 93; DB 2; Length 432;
Best Local Similarity 100.0%; Pred. No. 2.3e-06;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 QQAQAQVNTIQGVSAQT 20
DB 48 QQAQAQVNTIQGVSAQT 67

RESULT 7
Q5M6K4_STRT2
ID Q5M6K4_STRT2 PRELIMINARY; PRT; 474 AA.
AC Q5M6K4;
DT 01-FEB-2005 (TrEMBLrel. 29, Created)
DT 01-FEB-2005 (TrEMBLrel. 29, Last sequence update)
DT 13-SEP-2005 (TrEMBLrel. 31, Last annotation update)
DE Glucan binding protein (PcsB).
GN Name=pcsb; OrderedLocustNames=stu0022;
OS Streptococcus thermophilus (strain ATCC BAA-250 / LMG 18311).
OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
OC Streptococcus.
OX NCBI_TaxID=264199;
RN [1]
RP NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
RX PubMed=15543133; DOI=10.1038/nbt1034;
RA Bolotin A., Quinquis B., Renault P., Sorokin A., Ehrlich S.D.,
RA Kulakauskas S., Lapidus A., Goldsman E., Mazur M., Pusch G.D.,
RA Fonstein M., Overbeek R., Kyprides N., Purnelle B., Prozzi D.,
RA Ngui K., Masuy D., Hancy F., Burteau S., Boutry M., Delcour J.,
RA Goffeau A., Hols P.;
RT "Complete sequence and comparative genome analysis of the dairy
RT bacterium Streptococcus thermophilus.";
RL Nat. Biotechnol. 22:1554-1558(2004).
RN [2]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=LMG 18311;
RA Borges F., Layec S., Thibessard A., Fernandez A., Gintz B., Hols P.,
RA Decaris B., Leblond-Bourget N.;
RT "cse, a Chimeric and Variable Gene, Encodes an Extracellular Protein
RT Involved in Cellular Segregation in Streptococcus thermophilus.";
RL J. Bacteriol. 187:2737-2746(2005).
DR EMBL; CP000023; AAV59752.1; -; Genomic_DNA.
DR EMBL; AY730643; AAW82375.1; -; Genomic_DNA.
DR InterPro; IPR007921; CHAP.
DR Pfam; PF05257; CHAP; 1.
DR PRINTS; PR01852; SIBAPROTEIN.
DR PROSITE; PS50911; CHAP; 1.
KW Complete proteome.
SQ SEQUENCE 474 AA; 48142 MW; 1BAC6A9E0A0A200D CRC64;

Query Match 73.1%; Score 68; DB 2; Length 474;
Best Local Similarity 75.0%; Pred. No. 0.027;
Matches 15; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 1 QQAQAQVNTIQGVSAQT 20

```

```

RC STRAIN=15JP2;
RX MEDLINE=21481971; PubMed=11598068;
RX DOI=10.1128/JAI.69.11.6931-6941.2001;
RA Mattos-Graner R.O., Jin S., King W.F., Chen T., Smith D.J.,
RA Duncan M.J.;
RT "Cloning of the Streptococcus mutans gene encoding glucan binding
RT protein B and analysis of genetic diversity and protein production in
RT clinical isolates.";
RL Infect. Immun. 69:6931-6941(2001).
DR EMBL; AY046412; AAK94502.1; -; Genomic_DNA.
DR InterPro; IPR007921; CHAP.
DR Pfam; PF05257; CHAP; 1.
DR PRINTS; PR01852; SIBAPROTEIN.
DR PROSITE; PS50911; CHAP; 1.
SQ SEQUENCE 432 AA; 44648 MW; E769B2504AEE50E9 CRC64;

Query Match 100.0%; Score 93; DB 2; Length 432;
Best Local Similarity 100.0%; Pred. No. 2.3e-06;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 QQAQAQVNTIQGVSAQT 20
DB 48 QQAQAQVNTIQGVSAQT 67

RESULT 7
Q5M6K4_STRT2
ID Q5M6K4_STRT2 PRELIMINARY; PRT; 474 AA.
AC Q5M6K4;
DT 01-FEB-2005 (TrEMBLrel. 29, Created)
DT 01-FEB-2005 (TrEMBLrel. 29, Last sequence update)
DT 13-SEP-2005 (TrEMBLrel. 31, Last annotation update)
DE Glucan binding protein (PcsB).
GN Name=pcsb; OrderedLocustNames=stu0022;
OS Streptococcus thermophilus (strain ATCC BAA-250 / LMG 18311).
OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
OC Streptococcus.
OX NCBI_TaxID=264199;
RN [1]
RP NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
RX PubMed=15543133; DOI=10.1038/nbt1034;
RA Bolotin A., Quinquis B., Renault P., Sorokin A., Ehrlich S.D.,
RA Kulakauskas S., Lapidus A., Goldsman E., Mazur M., Pusch G.D.,
RA Fonstein M., Overbeek R., Kyprides N., Purnelle B., Prozzi D.,
RA Ngui K., Masuy D., Hancy F., Burteau S., Boutry M., Delcour J.,
RA Goffeau A., Hols P.;
RT "Complete sequence and comparative genome analysis of the dairy
RT bacterium Streptococcus thermophilus.";
RL Nat. Biotechnol. 22:1554-1558(2004).
RN [2]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=LMG 18311;
RA Borges F., Layec S., Thibessard A., Fernandez A., Gintz B., Hols P.,
RA Decaris B., Leblond-Bourget N.;
RT "cse, a Chimeric and Variable Gene, Encodes an Extracellular Protein
RT Involved in Cellular Segregation in Streptococcus thermophilus.";
RL J. Bacteriol. 187:2737-2746(2005).
DR EMBL; CP000023; AAV59752.1; -; Genomic_DNA.
DR EMBL; AY730643; AAW82375.1; -; Genomic_DNA.
DR InterPro; IPR007921; CHAP.
DR Pfam; PF05257; CHAP; 1.
DR PRINTS; PR01852; SIBAPROTEIN.
DR PROSITE; PS50911; CHAP; 1.
KW Complete proteome.
SQ SEQUENCE 474 AA; 48142 MW; 1BAC6A9E0A0A200D CRC64;

Query Match 73.1%; Score 68; DB 2; Length 474;
Best Local Similarity 75.0%; Pred. No. 0.027;
Matches 15; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 1 QQAQAQVNTIQGVSAQT 20

```

Db 45 QEAQAQVATIQSVSTLRT 64
|:|||||:|||||:|

RESULT 8
QSM212 STRT1 PRELIMINARY; PRT; 485 AA.
AC QSM212_29 (TREMBLrel. 29, Created)
DT 01-FEB-2005 (TREMBLrel. 29, Last sequence update)
DT 01-FEB-2005 (TREMBLrel. 29, Last annotation update)
DE Glucan binding protein.
GN Name=pcsb; OrderedLocusNames=st0022;
OS Streptococcus thermophilus (strain CNRZ 1066).
OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
OC Streptococcus.
OX NCBI_TaxID=299768;
RN [1]
RP NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
RX PubMed=15543133; DOI=10.1038/nbt1034;
RA Bolotin A., Quinquis B., Renault P., Sorokin A., Ehrlich S.D.,
RA Kulakauskas S., Lapidus A., Goldstein E., Mazur M., Pusch G.D.,
RA Fonstein M., Overbeek R., Kyprides N., Purnelle B., Prozzi D.,
RA Ngui K., Masuy D., Hancy F., Burteau S., Boutry M., Delcours J.,
RA Goffeau A., Hols P.;
RT "Complete sequence and comparative genome analysis of the dairy
bacterium Streptococcus thermophilus";
RL Nat. Biotechnol. 22:1554-1558 (2004).
DR EMBL; CP000024; AAV61641.1; -; Genomic_DNA.
DR InterPro; IPR007921; CHAP.
DR InterPro; IPR009148; Siba.
DR Pfam; PF05257; CHAP; 1.
DR PRINTS; PR01852; SIBAPROTEIN.
DR PROSITE; PS50911; CHAP; 1.
KW Complete proteome.
SQ SEQUENCE 485 AA; 49165 MW; 8665E4B1FC6CF7CA CRC64;

Query Match 73.1%; Score 68; DB 2; Length 485;
Best Local Similarity 75.0%; Pred. No. 0.027;
Matches 15; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 1 QQAQAQVNTIQGVSAIQT 20
|:|||||:|||||:|

Db 45 QEAQAQVATIQSVSTLRT 64
|:|||||:|||||:|

RESULT 9
QXEL1 STRP6 PRELIMINARY; PRT; 398 AA.
AC QXEL1_28 (TREMBLrel. 28, Created)
DT 25-OCT-2004 (TREMBLrel. 28, Last sequence update)
DT 25-OCT-2004 (TREMBLrel. 28, Last annotation update)
DE Secreted protein.
GN OrderedLocusNames=M6_Spy0017;
OS Streptococcus pyogenes (serotype M6).
OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
OC Streptococcus.
OX NCBI_TaxID=301450;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=MGAS10394;
RX PubMed=15272401; DOI=10.1086/422697;
RA Banks D.J., Porcella S.F., Barbian K.D., Beres S.B., Phillips L.E.,
RA Voyich J.M., DeLeo F.R., Martin J.M., Somerville G.A., Musser J.M.;
RT "Progress toward characterization of the group A Streptococcus
metagenome: complete genome sequence of a macrolide-resistant serotype
M6 strain";
RL J. Infect. Dis. 190:727-738 (2004).
DR EMBL; CP000003; AAT86152.1; -; Genomic_DNA.
DR InterPro; IPR007921; CHAP.
DR InterPro; IPR009148; Siba.
DR Pfam; PF05257; CHAP; 1.

DR PRINTS; PR01852; SIBAPROTEIN.
DR PROSITE; PS50911; CHAP; 1.
KW Complete proteome.
SQ SEQUENCE 398 AA; 42028 MW; 5DCEDA78CB863B60 CRC64;

Query Match 62.4%; Score 58; DB 2; Length 398;
Best Local Similarity 60.0%; Pred. No. 0.92;
Matches 12; Conservative 5; Mismatches 3; Indels 0; Gaps 0;

QY 1 QQAQAQVNTIQGVSAIQT 20
|:|||||:|||||:|

Db 45 QKAQAQVSAIQAQVSSIQS 64
|:|||||:|||||:|

RESULT 10
Q9A128 STRPY PRELIMINARY; PRT; 398 AA.
AC Q9A128_17 (TREMBLrel. 17, Created)
DT 01-JUN-2001 (TREMBLrel. 17, Last sequence update)
DT 10-MAY-2005 (TREMBLrel. 30, Last annotation update)
DE Secreted protein Siba precursor.
GN OrderedLocusNames=SPY0019;
OS Streptococcus pyogenes.
OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
OC Streptococcus.
OX NCBI_TaxID=1314;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=SF370 / ATCC 700294 / Serotype M1;
RX MEDLINE=21192684; PubMed=11296296; DOI=10.1073/pnas.071559398;
RA Perretti J.J., McShan W.M., Ajdic D.J., Savic D.J., Savic G., Lyon K.,
RA Primeaux C., Sezate S., Suvorov A.N., Kenton S., Lai H.S., Lin S.P.,
RA Qian X., Jia H.G., Najjar F.Z., Ren Q., Zhu H., Song L., White J.,
RA Yuan X., Clifton S.W., Roe B.A., McLaughlin R.E.;
RT "Complete genome sequence of an M1 strain of Streptococcus pyogenes";
RL Proc. Natl. Acad. Sci. U.S.A. 98:4658-4663 (2001).
RN [2]
RP NUCLEOTIDE SEQUENCE.
RC Fagan P.K., Reinscheid D., Gottschalk B., Chatwal G.S.;
RT "Identification and characterization of a novel secreted protein from
group A streptococcus";
RL Submitted (NOV-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AE006474; AAK33158.1; -; Genomic_DNA.
DR EMBL; AF319999; AAL73135.1; -; Genomic_DNA.
DR InterPro; IPR007921; CHAP.
DR InterPro; IPR009148; Siba.
DR Pfam; PF05257; CHAP; 1.
DR PRINTS; PR01852; SIBAPROTEIN.
DR PROSITE; PS50911; CHAP; 1.
KW Complete proteome; Signal.
FT SIGNAL 1 23 Potential.
SQ SEQUENCE 398 AA; 41899 MW; 28A9B3F7195E969B CRC64;

Query Match 62.4%; Score 58; DB 2; Length 398;
Best Local Similarity 60.0%; Pred. No. 0.92;
Matches 12; Conservative 5; Mismatches 3; Indels 0; Gaps 0;

QY 1 QQAQAQVNTIQGVSAIQT 20
|:|||||:|||||:|

Db 45 QKAQAQVSAIQAQVSSIQS 64
|:|||||:|||||:|

RESULT 11
Q7CNQ7 STRP8 PRELIMINARY; PRT; 398 AA.
AC Q7CNQ7_27 (TREMBLrel. 27, Created)
DT 05-JUL-2004 (TREMBLrel. 27, Last sequence update)
DT 05-JUL-2004 (TREMBLrel. 27, Last annotation update)
DE Putative secreted protein.
GN OrderedLocusNames=spym18_0020;
OS Streptococcus pyogenes (serotype M18).

OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
 OC Streptococcus
 OX NCBI_TaxID=301451;
 RN [1]
 RP NUCLEOTIDE SEQUENCE.
 RC STRAIN=MGAS8232 / Serotype M18;
 RX MEDLINE=21927593; PubMed=11917108; DOI=10.1073/pnas.062526099;
 RA Smoot J.C., Barbican K.D., Van Gompel J.J., Smoot L.M., Chaussee M.S.,
 RA Sylva G.L., Sturdevant D.E., Ricklefs S.M., Porcella S.F.,
 RA Parkins L.D., Beres S.B., Campbell D.S., Smith T.M., Zhang Q.,
 RA Kapur V., Daly J.A., Veasy L.G., Musser J.M.;
 RT "Genome sequence and comparative microarray analysis of serotype M18
 RT group A Streptococcus strains associated with acute rheumatic fever
 RT outbreaks";
 RL Proc. Natl. Acad. Sci. U.S.A. 99:4668-4673 (2002).
 DR EMBL; AE009955; AAL96849.1; -; Genomic_DNA.
 DR InterPro; IPR007921; CHAP.
 DR Pfam; PF05257; CHAP; 1.
 DR PRINTS; PR01852; SIBAPROTEIN.
 DR PROSITE; PS50911; CHAP; 1.
 KW Complete proteome.
 SQ SEQUENCE 398 AA; 42028 MW; 5DCEDA78CB863B60 CRC64;
 Query Match 62.4%; Score 58; DB 2; Length 398;
 Best Local Similarity 60.0%; Pred. No. 0.92;
 Matches 12; Conservative 5; Mismatches 3; Indels 0; Gaps 0;
 OY 1 QQAQAQVNTTIQGVSAIQT 20
 DB 45 QKAQAQVSAIQAQVSSLSQS 64
 RESULT 12
 ID QBP318 STRP3 PRELIMINARY; PRT; 398 AA.
 AC QBP318_Q7CFL7;
 DT 01-OCT-2002 (TrEMBLrel. 22, Created)
 DT 01-OCT-2002 (TrEMBLrel. 22, Last sequence update)
 DT 13-SEP-2005 (TrEMBLrel. 31, Last annotation update)
 DE Putative secreted protein.
 GN OrderedLocusNames=SPS0015, SPM3_0014;
 OS Streptococcus pyogenes (serotype M3).
 OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
 OC Streptococcus.
 OX NCBI_TaxID=301448;
 RN [1]
 RP NUCLEOTIDE SEQUENCE.
 RC STRAIN=SSI-1 / Serotype M3;
 RX MEDLINE=22683278; PubMed=12799345; DOI=10.1101/gr.1096703;
 RA Nakagawa I., Kurokawa K., Yamashita A., Nakata M., Tomiyasu Y.,
 RA Okahashi N., Kawabata S., Yamazaki K., Shiba T., Yasunaga T.,
 RA Hayashi H., Hattori M., Hamada S.;
 RT "Genome sequence of an M3 strain of Streptococcus pyogenes reveals a
 RT large-scale genomic rearrangement in invasive strains and new insights
 RT into phage evolution.";
 RL Genome Res. 13:1042-1055 (2003).
 RN [2]
 RP NUCLEOTIDE SEQUENCE.
 RC STRAIN=MGAS315 / Serotype M3;
 RX MEDLINE=22133808; PubMed=12122206; DOI=10.1073/pnas.152298499;
 RA Beres S.B., Sylva G.L., Barbican K.D., Lei B., Hoff J.S.,
 RA Mammarella N.D., Liu M.-Y., Smoot J.C., Porcella S.F., Parkins L.D.,
 RA Campbell D.S., Smith T.M., McCormick J.K., Leung D.Y.M.,
 RA Schlievert P.M., Musser J.M.;
 RT "Genome sequence of a serotype M3 strain of group A Streptococcus;
 RT phage-encoded toxins, the high-virulence phenotype, and clone
 RT emergence";
 RL Proc. Natl. Acad. Sci. U.S.A. 99:10078-10083 (2002).
 DR EMBL; BA000034; BAC63110.1; -; Genomic_DNA.
 DR EMBL; AE014136; AAM78621.1; -; Genomic_DNA.
 DR InterPro; IPR007921; CHAP.
 DR InterPro; IPR009148; SIDA.

DR Pfam; PF05257; CHAP; 1.
 DR PRINTS; PR01852; SIBAPROTEIN.
 DR PROSITE; PS50911; CHAP; 1.
 KW Complete proteome.
 SQ SEQUENCE 398 AA; 42028 MW; 5DCEDA78CB863B60 CRC64;
 Query Match 62.4%; Score 58; DB 2; Length 398;
 Best Local Similarity 60.0%; Pred. No. 0.92;
 Matches 12; Conservative 5; Mismatches 3; Indels 0; Gaps 0;
 OY 1 QQAQAQVNTTIQGVSAIQT 20
 DB 45 QKAQAQVSAIQAQVSSLSQS 64
 RESULT 13
 ID Q8DMV4 STRR6 PRELIMINARY; PRT; 392 AA.
 AC Q8DMV4;
 DT 01-MAR-2003 (TrEMBLrel. 23, Created)
 DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
 DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
 DE General stress protein GSP-781.
 GN Name=gsp-781; OrderedLocusNames=spr2021;
 OS Streptococcus pneumoniae (strain ATCC BAA-255 / R6).
 OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
 OC Streptococcus.
 OX NCBI_TaxID=171101;
 RN [1]
 RP NUCLEOTIDE SEQUENCE.
 RX MEDLINE=21429245; PubMed=11544234;
 RX DOI=10.1128/JB.183.19.5709-5717.2001;
 RA Hoskins J., Alborn W.E. Jr., Arnold J., Blaszcak L.C., Burgett S.,
 RA DeHoff B.S., Estrem S.T., Fritz L., Fu D.-J., Fuller W., Geringer C.,
 RA Gilmore R., Glass J.S., Khoja H., Kraft A.R., Lagace R.E.,
 RA LeBlanc D.J., Lee L.N., Lefkowitz E.J., Lu J., Matsushima P.,
 RA Mahren S.M., McHenry M., McLeaster K., Mundy C.W., Nicas T.I.,
 RA Norris F.H., O'Gara M., Peery R.B., Robertson G.T., Rockey P.,
 RA Sun P.-M., Winkler M.E., Yang Y., Young-Bellido M., Zhao G.,
 RA Zook C.A., Baltz R.H., Jaskunas S.R., Rostek P.R. Jr., Skatrud P.L.,
 RA Glass J.I.;
 RT "Genome of the bacterium Streptococcus pneumoniae strain R6.";
 RL J. Bacteriol. 183:5709-5717 (2001).
 DR EMBL; AE008566; AAL00823.1; -; Genomic_DNA.
 DR PIR; B98124; B98124.
 DR PIR; G95258; G95258.
 DR InterPro; IPR007921; CHAP.
 DR InterPro; IPR009148; SIDA.
 DR Pfam; PF05257; CHAP; 1.
 DR PRINTS; PR01852; SIBAPROTEIN.
 DR PROSITE; PS50911; CHAP; 1.
 KW Complete proteome.
 SQ SEQUENCE 392 AA; 41697 MW; B5A5A860FAEA77DD CRC64;
 Query Match 61.3%; Score 57; DB 2; Length 392;
 Best Local Similarity 68.4%; Pred. No. 1.3;
 Matches 13; Conservative 2; Mismatches 4; Indels 0; Gaps 0;
 OY 1 QQAQAQVNTTIQGVSAIQ 19
 DB 48 QQAQAQVDTIQEQVSAIQ 66
 RESULT 14
 ID Q97N55 STRPN PRELIMINARY; PRT; 392 AA.
 AC Q97N55;
 DT 01-OCT-2001 (TrEMBLrel. 18, Created)
 DT 01-OCT-2001 (TrEMBLrel. 18, Last sequence update)
 DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
 DE Secreted 45 kd protein.
 GN OrderedLocusNames=SP2216;
 OS Streptococcus pneumoniae.

OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
OC Streptococcus.
OX NCBI_TaxID=1313;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=ATCC BAA-334 / TIGR4;
RX MEDLINE=21357209; PubMed=11463916; DOI=10.1126/science.1061217;
RA Tettelin N., Nelson K.E., Paulsen I.T., Eisen J.A., Read T.D.,
RA Peterson S.N., Heidelberg J.F., DeBoy R.T., Haft D.H., Dodson R.J.,
RA Durkin A.S., Gwinn M.L., Kolonay J.F., Nelson W.C., Peterson J.D.,
RA Umayam L.A., White O., Salzberg S.L., Lewis M.R., Radune D.,
RA Holtzapple E.K., Khouri H.M., Wolf A.M., Utterback T.R., Hansen C.L.,
RA McDonald L.A., Feldblyum T.V., Angiuoli S.V., Dickinson T.,
RA Hickey E.K., Holt I.B., Loftus B.J., Yang F., Smith H.O., Venter J.C.,
RA Dougherty B.A., Morrison D.A., Hollingshead S.K., Fraser C.M.;
RT "Complete genome sequence of a virulent isolate of Streptococcus
RT pneumoniae.";
RL Science 293:498-506(2001).
DR EMBL; AE007509; AAK76264.1; -; Genomic_DNA.
DR FIR; B98124; B98124.
DR PIR; G95258; G95258.
DR TIGR; SP2216; -.
DR InterPro; IPR007921; CHAP.
DR InterPro; IPR009148; Siba.
DR Pfam; PF05257; CHAP; 1.
DR PRINTS; PRO1852; SIHAPROTEIN.
DR PROSITE; PS50911; CHAP; 1.
KW Complete proteome.
SQ SEQUENCE 392 AA; 41697 MW; B5A5A860FAEA77DD CRC64;

Query Match 61.3%; Score 57; DB 2; Length 392;
Best Local Similarity 68.4%; Pred. No. 1.3;
Matches 13; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

Qy 1 QQAQAQAVNTTIOGVSAIQ 19
Db 48 QQAQAQAVNTTIOGVSAIQ 66

RESULT 15
P54 ENTFC
ID_P54_ENTFC STANDARD; PRT; 516 AA.
AC P13692;
DT 01-JAN-1990 (Rel. 13, Created)
DT 01-NOV-1991 (Rel. 20, Last sequence update)
DT 13-SEP-2005 (Rel. 48, Last annotation update)
DE P54 protein precursor.
OS Enterococcus faecium (Streptococcus faecium).
OC Bacteria; Firmicutes; Lactobacillales; Enterococcaceae; Enterococcus.
OX NCBI_TaxID=1352;
RN [1]
RP NUCLEOTIDE SEQUENCE [GENOMIC DNA].
RX MEDLINE=89385998; PubMed=2780297;
RA Fuerst P., Moesch H.-U., Solioz M.;
RT "A protein of unusual composition from Enterococcus faecium.";
RL Nucleic Acids Res. 17:6724-6724(1989).
CC -!- SUBCELLULAR LOCATION: Attached to the cell wall.
CC -!- SIMILARITY: Belongs to the peptidase C40 family.
CC -----
CC This Swiss-Prot entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use as long as its content is in no way modified and this statement is not
CC removed.
CC -----
DR EMBL; X16421; CAA34442.1; ALT_INIT; Genomic_DNA.
DR PIR; S05542; S05542.
DR InterPro; IPR000064; NLP_P60.
DR Pfam; PF00877; NLPFC_P60; 1.
KW Cell wall; Signal.
FT SIGNAL 1 27 Potential.
FT CHAIN 28 516 P54 protein.
FT ACT_SITE 429 429 Potential.

SQ SEQUENCE 516 AA; 54596 MW; 402ECAAA39846D26 CRC64;

Query Match 61.3%; Score 57; DB 1; Length 516;
Best Local Similarity 50.0%; Pred. No. 1.7;
Matches 10; Conservative 6; Mismatches 4; Indels 0; Gaps 0;

Qy 1 QQAQAQAVNTTIOGVSAIQ 20
Db 48 QQAQAQAVNTTIOGVSAIQ 67

Search completed: January 25, 2006, 19:13:40
Job time : 71.3 secs


```

; TYPE: PRT
; ORGANISM: Streptococcus pneumoniae
US-09-583-110-4374

Query Match      61.3%; Score 57; DB 2; Length 392;
Best Local Similarity 68.4%; Pred. No. 0.15;
Matches 13; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

QY 1 QQAQAQVNTIQGVSAIQ 19
   |||||:|||||:|
Db 48 QQAQAQVNTIQGVSAIQ 66

RESULT 3
US-09-107-433-3230
; Sequence 3230, Application US/09107433
; Patent No. 6800744
; GENERAL INFORMATION:
; APPLICANT: Lynn A Doucette-Stamm and David Bush
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID
; SEQUENCES RELATING TO STREPTOCOCCUS PNEUMONIAE FOR DIAGNOSTIC
; THERAPEUTICS
; NUMBER OF SEQUENCES: 5206
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: GENOME THERAPEUTICS CORPORATION
; STREET: 100 Beaver Street
; CITY: Waltham
; STATE: Massachusetts
; COUNTRY: USA
; ZIP: 02354
; COMPUTER READABLE FORM:
; MEDIUM TYPE: CD-ROM ISO9660
; COMPUTER: <Unknown>
; OPERATING SYSTEM: <Unknown>
; SOFTWARE: <Unknown>
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/107,433
; FILING DATE: 30-Jun-1998
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 60/ 085131
; FILING DATE: May 12, 1998
; APPLICATION NUMBER: 60/051553
; FILING DATE: July 2, 1997
; ATTORNEY/AGENT INFORMATION:
; NAME: Ariniello, Pamela Deneke
; REGISTRATION NUMBER: 40,489
; REFERENCE/DOCKET NUMBER: GTC-011
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (781)893-5007
; TELEFAX: (781)893-8277
; INFORMATION FOR SEQ ID NO: 3230:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 399 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; HYPOTHETICAL: YES
; ORIGINAL SOURCE:
; ORGANISM: Streptococcus pneumoniae
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (B) LOCATION 1....399
; SEQUENCE DESCRIPTION: SEQ ID NO: 3230:
US-09-107-433-3230

Query Match      61.3%; Score 57; DB 2; Length 399;
Best Local Similarity 68.4%; Pred. No. 0.15;
Matches 13; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

QY 1 QQAQAQVNTIQGVSAIQ 19
   |||||:|||||:|
Db 55 QQAQAQVNTIQGVSAIQ 73

```

```

RESULT 4
US-09-562-737-45
; Sequence 45, Application US/09562737
; Patent No. 6428967
; GENERAL INFORMATION:
; APPLICANT: Herz, Joachim
; TITLE OF INVENTION: LDL Receptor Signaling Pathways
; FILE REFERENCE: UTSW0708
; CURRENT APPLICATION NUMBER: US/09/562,737
; CURRENT FILING DATE: 2000-05-01
; NUMBER OF SEQ ID NOS: 132
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 45
; LENGTH: 1024
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Synthetic
; OTHER INFORMATION: Sequence
US-09-562-737-45

Query Match      53.8%; Score 50; DB 2; Length 1024;
Best Local Similarity 70.6%; Pred. No. 6.9;
Matches 12; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

QY 3 AAQAQVNTIQGVSAIQ 19
   |||||:|||||:|
Db 1003 AATASVPTIQQVSAMQ 1019

RESULT 5
US-09-562-737-41
; Sequence 41, Application US/09562737
; Patent No. 6428967
; GENERAL INFORMATION:
; APPLICANT: Herz, Joachim
; TITLE OF INVENTION: LDL Receptor Signaling Pathways
; FILE REFERENCE: UTSW0708
; CURRENT APPLICATION NUMBER: US/09/562,737
; CURRENT FILING DATE: 2000-05-01
; NUMBER OF SEQ ID NOS: 132
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 41
; LENGTH: 1024
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Synthetic
; OTHER INFORMATION: Sequence
US-09-562-737-41

Query Match      51.6%; Score 48; DB 2; Length 1024;
Best Local Similarity 64.7%; Pred. No. 15;
Matches 11; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

QY 3 AAQAQVNTIQGVSAIQ 19
   |||||:|||||:|
Db 1003 AAKASVPTIQQASAMQ 1019

RESULT 6
US-09-562-737-47
; Sequence 47, Application US/09562737
; Patent No. 6428967
; GENERAL INFORMATION:
; APPLICANT: Herz, Joachim
; TITLE OF INVENTION: LDL Receptor Signaling Pathways
; FILE REFERENCE: UTSW0708
; CURRENT APPLICATION NUMBER: US/09/562,737

```

; CURRENT FILING DATE: 2000-05-01
 ; NUMBER OF SEQ ID NOS: 132
 ; SOFTWARE: PatentIn Ver. 2.1
 ; SEQ ID NO 47
 ; TYPE: PRT
 ; ORGANISM: Artificial Sequence
 ; FEATURE:
 ; OTHER INFORMATION: Description of Artificial Sequence: Synthetic
 ; OTHER INFORMATION: Sequence
 US-09-562-737-47

Query Match 51.6%; Score 48; DB 2; Length 1024;
 Best Local Similarity 64.7%; Pred. No. 15;
 Matches 11; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

QY 3 AQAQVNTIQGVSAIQ 19
 |||:|||||:|
 Db 1003 AKAQVPTIQQASRMQ 1019

RESULT 7
 US-08-186-222-2
 ; Sequence 2, Application US/08186222
 ; Patent No. 5559007
 ; GENERAL INFORMATION:
 ; APPLICANT: Suri, Bruno
 ; APPLICANT: Schmitz, Albert
 ; TITLE OF INVENTION: Bacterial Vectors
 ; NUMBER OF SEQUENCES: 4
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: CIBA-GEIGY Corporation
 ; STREET: 7 Skyline Drive
 ; CITY: Hawthorne
 ; STATE: New York
 ; COUNTRY: USA
 ; ZIP: 10532
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Floppy disk
 ; COMPUTER: IBM PC compatible
 ; OPERATING SYSTEM: PC-DOS/MS-DOS
 ; SOFTWARE: PatentIn Release #1.0, Version #1.25
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/08/186,222
 ; FILING DATE:
 ; CLASSIFICATION: 435
 ; PRIOR APPLICATION NUMBER: US 07/672,205
 ; FILING DATE: 19-MAR-1991
 ; APPLICATION NUMBER: GB 9006400.7
 ; FILING DATE: 22-MAR-1990
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: Villamizar, JoAnn
 ; REGISTRATION NUMBER: 30,598
 ; REFERENCE/DOCKET NUMBER: 4-17994/A
 ; TELECOMMUNICATION INFORMATION:
 ; TELEPHONE: (914)785-7121
 ; TELEFAX: (914)347-5769
 ; INFORMATION FOR SEQ ID NO: 2:
 ; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 461 amino acids
 ; TYPE: amino acid
 ; TOPOLOGY: linear
 ; MOLECULE TYPE: protein
 US-08-186-222-2

Query Match 49.5%; Score 46; DB 1; Length 461;
 Best Local Similarity 56.2%; Pred. No. 13;
 Matches 9; Conservative 5; Mismatches 2; Indels 0; Gaps 0;

QY 4 AQAQVNTIQGVSAIQ 19
 |||:|||||:|
 Db 50 AQAQVDSLQSKVDSLQ 65

RESULT 8
 US-08-816-693A-51
 ; Sequence 51, Application US/08816693A
 ; Patent No. 5874241
 ; GENERAL INFORMATION:
 ; APPLICANT: Takahashi, Joseph S
 ; APPLICANT: Turek, Fred W
 ; APPLICANT: Pinto, Lawrence H
 ; TITLE OF INVENTION: Clock Gene and Gene Product
 ; NUMBER OF SEQUENCES: 53
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: Dressler, Rocky, Milnamow & Katz
 ; STREET: Two Prudential Plaza, Suite 4700
 ; CITY: Chicago
 ; STATE: Illinois
 ; COUNTRY: USA
 ; ZIP: 60601
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Floppy disk
 ; COMPUTER: IBM PC compatible
 ; OPERATING SYSTEM: PC-DOS/MS-DOS
 ; SOFTWARE: PatentIn Release #1.0, Version #1.30
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/08/816,693A
 ; FILING DATE:
 ; CLASSIFICATION: 435
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: No. 5874241chup, Thomas E
 ; REGISTRATION NUMBER: 33,268
 ; TELECOMMUNICATION INFORMATION:
 ; TELEPHONE: 312-616-5400
 ; TELEFAX: 312-616-5460
 ; INFORMATION FOR SEQ ID NO: 51:
 ; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 747 amino acids
 ; TYPE: amino acid
 ; STRANDEDNESS: single
 ; TOPOLOGY: linear
 ; MOLECULE TYPE: peptide
 US-08-816-693A-51

Query Match 49.5%; Score 46; DB 1; Length 747;
 Best Local Similarity 50.0%; Pred. No. 22;
 Matches 10; Conservative 4; Mismatches 6; Indels 0; Gaps 0;

QY 1 QQAQAQVNTIQGVSAIQT 20
 |||:|||||:|
 Db 545 QRAAQPLVQLQGQISTQVT 564

RESULT 9
 US-08-885-291-51
 ; Sequence 51, Application US/08885291A
 ; Patent No. 6057125
 ; GENERAL INFORMATION:
 ; APPLICANT: Takahashi, Joseph S.
 ; APPLICANT: Turek, Fred W.
 ; APPLICANT: Pinto, Lawrence H.
 ; TITLE OF INVENTION: CLOCK GENE AND GENE PRODUCT
 ; FILE REFERENCE: 0290-5
 ; CURRENT APPLICATION NUMBER: US/08/885,291A
 ; CURRENT FILING DATE: 1997-06-30
 ; EARLIER APPLICATION NUMBER: 08/816,693
 ; EARLIER FILING DATE: 1997-03-13
 ; NUMBER OF SEQ ID NOS: 55
 ; SOFTWARE: PatentIn Ver. 2.0
 ; SEQ ID NO 51
 ; LENGTH: 747
 ; TYPE: PRT
 ; ORGANISM: Mus musculus
 US-08-885-291-51

```

Query Match          49.5%; Score 46; DB 2; Length 747;
Best Local Similarity 50.0%; Pred. No. 22;
Matches 10; Conservative 4; Mismatches 6; Indels 0; Gaps 0;

QY 1 QQAQAQVNTTIGQVSALQT 20
   ||||| | : ||||| |
Db 545 QRAAQPQLVQLGQISTQVT 564

RESULT 10
US-09-496-672-51
; Sequence 51, Application US/09496672
; Patent No. 6291429
; GENERAL INFORMATION:
; APPLICANT: Takahashi, Joseph S.
; APPLICANT: Turek, Fred W.
; APPLICANT: Pinto, Lawrence H.
; TITLE OF INVENTION: CLOCK GENE AND GENE PRODUCT
; FILE REFERENCE: 0290-5
; CURRENT APPLICATION NUMBER: US/09/496,672
; CURRENT FILING DATE: 2000-02-03
; PRIOR APPLICATION NUMBER: 08/885,291
; PRIOR FILING DATE: 1997-06-30
; PRIOR APPLICATION NUMBER: 08/816,693
; PRIOR FILING DATE: 1997-03-13
; NUMBER OF SEQ ID NOS: 55
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 51
; LENGTH: 747
; TYPE: PRT
; ORGANISM: Mus musculus
US-09-496-672-51

Query Match          49.5%; Score 46; DB 2; Length 747;
Best Local Similarity 50.0%; Pred. No. 22;
Matches 10; Conservative 4; Mismatches 6; Indels 0; Gaps 0;

QY 1 QQAQAQVNTTIGQVSALQT 20
   ||||| | : ||||| |
Db 545 QRAAQPQLVQLGQISTQVT 564

RESULT 11
US-09-107-532A-5095
; Sequence 5095, Application US/09107532A
; Patent No. 6583275
; GENERAL INFORMATION:
; APPLICANT: Lynn A Doucette-Stamm and David Bush
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO
; ENTEROCOCCUS FAECIUM FOR DIAGNOSTICS AND THERAPEUTICS
; NUMBER OF SEQUENCES: 7310
; CORRESPONDENCE ADDRESS:
; ADDRESSER: GENOME THERAPEUTICS CORPORATION
; STREET: 100 Beaver Street
; CITY: Waltham
; STATE: Massachusetts
; COUNTRY: USA
; ZIP: 02354
; COMPUTER READABLE FORM:
; MEDIUM TYPE: CD-ROM ISO9660
; COMPUTER: PC
; OPERATING SYSTEM: <Unknown>
; SOFTWARE: ASCII
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/107,532A
; FILING DATE: 30-Jun-1998
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 60/085,598
; FILING DATE: 14 May 1998
; APPLICATION NUMBER: 60/051571
; FILING DATE: July 2, 1997
; ATTORNEY/AGENT INFORMATION:

```

```

; NAME: Ariniello, Pamela Deneke
; REGISTRATION NUMBER: 40,489
; REFERENCE/DOCKET NUMBER: GTC-012
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (781)893-5007
; TELEFAX: (781)893-8277
; INFORMATION FOR SEQ ID NO: 5095:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 525 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; HYPOTHETICAL: YES
; ORIGINAL SOURCE:
; ORGANISM: Enterococcus faecium
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (B) LOCATION 1...525
; SEQUENCE DESCRIPTION: SEQ ID NO: 5095:
US-09-107-532A-5095

Query Match          48.9%; Score 45.5; DB 2; Length 525;
Best Local Similarity 47.6%; Pred. No. 18;
Matches 10; Conservative 6; Mismatches 4; Indels 1; Gaps 1;

QY 1 QQA-AQAQVNTTIGQVSALQT 20
   ||||| ||||| : ||||| :
Db 48 QQAQAQSIDALESQVSEINT 68

RESULT 12
US-09-562-737-43
; Sequence 43, Application US/09562737
; Patent No. 6428967
; GENERAL INFORMATION:
; APPLICANT: Herz, Joachim
; APPLICANT: Gotthardt, Michael
; TITLE OF INVENTION: LDL Receptor Signaling Pathways
; FILE REFERENCE: UTSW0708
; CURRENT APPLICATION NUMBER: US/09/562,737
; CURRENT FILING DATE: 2000-05-01
; NUMBER OF SEQ ID NOS: 132
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 43
; LENGTH: 1024
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Synthetic
; OTHER INFORMATION: Sequence
US-09-562-737-43

Query Match          48.4%; Score 45; DB 2; Length 1024;
Best Local Similarity 62.5%; Pred. No. 48;
Matches 10; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

QY 4 AQAQVNTTIGQVSALQ 19
   | : | | | | | | | : |
Db 1004 AKASVPTTQAQASAMQ 1019

RESULT 13
US-09-543-681A-5472
; Sequence 5472, Application US/09543681A
; Patent No. 6605709
; GENERAL INFORMATION:
; APPLICANT: GARY BRETON
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PROTEUS MIRABILIS
; FILE REFERENCE: 2709.1002-001
; CURRENT APPLICATION NUMBER: US/09/543,681A
; CURRENT FILING DATE: 2000-04-05
; PRIOR APPLICATION NUMBER: US 60/128,706

```

; PRIOR FILING DATE: 1999-04-09
; NUMBER OF SEQ ID NOS: 8344
; SEQ ID NO 5472
; LENGTH: 94
; TYPE: PRT
; ORGANISM: Proteus mirabilis
US-09-543-681A-5472

Query Match 47.3%; Score 44; DB 2; Length 94;
Best Local Similarity 45.0%; Pred. No. 4.2;
Matches 9; Conservative 6; Mismatches 5; Indels 0; Gaps 0;

QY 1 QQAQAQVNTIQVSA 20
DB 11 KKAALTRVRKIGQILALET 30

RESULT 14
US-09-949-016-10850
; Sequence 10850, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 10850
; LENGTH: 773
; TYPE: PRT
; ORGANISM: Human
US-09-949-016-10850

Query Match 46.8%; Score 43.5; DB 2; Length 773;
Best Local Similarity 55.6%; Pred. No. 61;
Matches 10; Conservative 4; Mismatches 3; Indels 1; Gaps 1;

QY 1 QQAQAQVNTIQVSA 18
DB 493 QQAQAQNI-TLQEQIEAI 509

RESULT 15
US-09-489-039A-10762
; Sequence 10762, Application US/09489039A
; Patent No. 6610836
; GENERAL INFORMATION:
; APPLICANT: Gary Breton et. al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO KLEBSIELLA
; TITLE OF INVENTION: PNEUMONIAE FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 2709.2004001
; CURRENT APPLICATION NUMBER: US/09/489,039A
; CURRENT FILING DATE: 2000-01-27
; PRIOR APPLICATION NUMBER: US 60/117,747
; PRIOR FILING DATE: 1999-01-29
; NUMBER OF SEQ ID NOS: 14342
; SEQ ID NO 10762
; LENGTH: 231
; TYPE: PRT
; ORGANISM: Klebsiella pneumoniae
US-09-489-039A-10762

Query Match 46.2%; Score 43; DB 2; Length 231;
Best Local Similarity 47.1%; Pred. No. 18;

Matches 8; Conservative 6; Mismatches 3; Indels 0; Gaps 0;
QY 1 QQAQAQVNTIQVSA 17
DB 141 EQAARSALNSLQAFSA 157
Search completed: January 25, 2006, 19:18:14
Job time : 16.6 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2006 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: January 25, 2006, 19:13:51 ; Search time 60 Seconds
(without alignments)
139.276 Million cell updates/sec

Title: US-10-797-821-8
Perfect score: 93
Sequence: 1 QQAQAQVNTIQGVSAALQT 20

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1867569 seqs, 417829326 residues

Total number of hits satisfying chosen parameters: 1867569

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published Applications AA Main:*

- 1: /cgn2_6/ptodata/1/pubpaa/US07_PUBCOMB.pep.*
- 2: /cgn2_6/ptodata/1/pubpaa/US08_PUBCOMB.pep.*
- 3: /cgn2_6/ptodata/1/pubpaa/US09_PUBCOMB.pep.*
- 4: /cgn2_6/ptodata/1/pubpaa/US10A_PUBCOMB.pep.*
- 5: /cgn2_6/ptodata/1/pubpaa/US10B_PUBCOMB.pep.*
- 6: /cgn2_6/ptodata/1/pubpaa/US11_PUBCOMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	93	100.0	20	4	US-10-383-930-8
2	93	100.0	20	5	US-10-797-821-8
3	93	100.0	431	4	US-10-383-930-29
4	93	100.0	431	4	US-10-383-930-30
5	93	100.0	431	4	US-10-383-930-33
6	93	100.0	431	5	US-10-797-821-29
7	93	100.0	431	5	US-10-797-821-30
8	93	100.0	431	5	US-10-797-821-33
9	93	100.0	432	4	US-10-383-930-31
10	93	100.0	432	4	US-10-383-930-32
11	93	100.0	432	5	US-10-797-821-31
12	93	100.0	432	5	US-10-797-821-32
13	75	80.6	20	4	US-10-383-930-9
14	75	80.6	20	5	US-10-797-821-9
15	58	62.4	398	5	US-10-474-792-600
16	57	61.3	210	4	US-10-154-251-67
17	57	61.3	392	5	US-10-472-928-4652
18	57	61.3	399	5	US-10-617-320-3230
19	51	54.8	20	4	US-10-383-930-2
20	51	54.8	20	5	US-10-797-821-2
21	50	53.8	1024	4	US-10-211-962-45
22	48	51.6	825	4	US-10-369-493-22501
23	48	51.6	1024	4	US-10-211-962-41
24	48	51.6	1024	4	US-10-211-962-47
25	48	51.6	2540	5	US-10-606-038-6
26	48	51.6	2541	4	US-10-177-293-470
27	48	51.6	2541	5	US-10-719-993-535

28	48	51.6	2541	5	US-10-719-993-536	Sequence 536, App
29	48	51.6	2541	5	US-10-287-436A-460	Sequence 460, App
30	48	51.6	2541	5	US-10-287-436A-1259	Sequence 1259, App
31	47	50.5	218	4	US-10-437-963-199647	Sequence 199647, App
32	47	50.5	378	5	US-10-820-155-28	Sequence 28, Appl
33	47	50.5	378	5	US-10-820-155-62	Sequence 62, Appl
34	46	49.5	355	4	US-10-029-386-33377	Sequence 33377, A
35	46	49.5	454	4	US-10-282-122A-78518	Sequence 78518, A
36	46	49.5	824	4	US-10-282-122A-59097	Sequence 59097, A
37	46	49.5	1379	6	US-11-097-143-261	Sequence 261, App
38	46	49.5	2427	4	US-10-369-493-6734	Sequence 6734, App
39	45.5	48.9	524	4	US-10-282-122A-57658	Sequence 57658, A
40	45	48.4	1024	4	US-10-211-962-43	Sequence 43, Appl
41	45	48.4	1465	6	US-11-097-143-20292	Sequence 20292, A
42	45	48.4	1645	4	US-10-263-929-176	Sequence 176, App
43	44	47.3	488	5	US-10-739-930-6862	Sequence 6862, App
44	44	47.3	502	4	US-10-437-963-125588	Sequence 125588, App
45	43.5	46.8	70	3	US-09-864-761-43799	Sequence 43799, A

ALIGNMENTS

RESULT 1

US-10-383-930-8
; Sequence 8, Application US/10383930
; Publication No. US20040127400A1
; GENERAL INFORMATION:
; APPLICANT: Smith, Daniel J
; APPLICANT: Taubman, Martin A
; TITLE OF INVENTION: Immunogenicity of Glucan Binding Protein
; FILE REFERENCE: 25669-018
; CURRENT APPLICATION NUMBER: US/10/383,930
; CURRENT FILING DATE: 2003-03-07
; PRIOR APPLICATION NUMBER: 60/402,483
; PRIOR FILING DATE: 2002-08-08
; PRIOR APPLICATION NUMBER: 60/363,209
; PRIOR FILING DATE: 2002-03-07
; NUMBER OF SEQ ID NOS: 41
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 8
; LENGTH: 20
; TYPE: PRT
; ORGANISM: Streptococcus mutans
US-10-383-930-8

Query Match	100.0%	Score 93;	DB 4;	Length 20;
Best Local Similarity	100.0%	Pred. No. 7.6e-08;	Mismatches 0;	Indels 0;
Matches	20;	Conservative 0;	Gaps 0;	
Qy	1	QQAQAQVNTIQGVSAALQT 20		
Db	1	QQAQAQVNTIQGVSAALQT 20		

RESULT 2

US-10-797-821-8
; Sequence 8, Application US/10797821
; Publication No. US20050031633A1
; GENERAL INFORMATION:
; APPLICANT: Smith, Daniel J.
; APPLICANT: Taubman, Martin A.
; TITLE OF INVENTION: Glucan Binding Protein and Glycosyltransferase Immunogens
; FILE REFERENCE: 25669-020
; CURRENT APPLICATION NUMBER: US/10/797,821
; CURRENT FILING DATE: 2004-03-09
; PRIOR APPLICATION NUMBER: 10/383,930
; PRIOR FILING DATE: 2003-03-07
; PRIOR APPLICATION NUMBER: 60/363,209
; PRIOR FILING DATE: 2002-03-07
; PRIOR APPLICATION NUMBER: 60/402,483
; PRIOR FILING DATE: 2002-08-08
; PRIOR APPLICATION NUMBER: 09/290,049

; PRIOR FILING DATE: 1999-04-12
; PRIOR APPLICATION NUMBER: 60/081,550
; PRIOR FILING DATE: 1998-04-13
; PRIOR APPLICATION NUMBER: 60/115,142
; PRIOR FILING DATE: 1999-01-08
; NUMBER OF SEQ ID NOS: 45
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 8
; LENGTH: 20
; TYPE: PRT
; ORGANISM: Artificial
; FEATURE:
; OTHER INFORMATION: GbpB peptide
US-10-797-821-8

Query Match 100.0%; Score 93; DB 5; Length 20;
Best Local Similarity 100.0%; Pred. No. 7.6e-08;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 QQAQAQVNTTIGQVSALQT 20
| | | | | | | | | | | | | | | | | | | | | |
Db 1 QQAQAQVNTTIGQVSALQT 20

RESULT 3
US-10-383-930-29
; Sequence 29, Application US/10383930
; Publication No. US20040127400A1
; GENERAL INFORMATION:
; APPLICANT: Smith, Daniel J
; APPLICANT: Taubman, Martin A
; TITLE OF INVENTION: Immunogenicity of Glucan Binding Protein
; FILE REFERENCE: 25669-018
; CURRENT APPLICATION NUMBER: US/10/383,930
; CURRENT FILING DATE: 2003-03-07
; PRIOR APPLICATION NUMBER: 60/402,483
; PRIOR FILING DATE: 2002-08-08
; PRIOR APPLICATION NUMBER: 60/363,209
; PRIOR FILING DATE: 2002-03-07
; NUMBER OF SEQ ID NOS: 41
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 29
; LENGTH: 431
; TYPE: PRT
; ORGANISM: Streptococcus mutans
US-10-383-930-29

Query Match 100.0%; Score 93; DB 4; Length 431;
Best Local Similarity 100.0%; Pred. No. 2.3e-06;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 QQAQAQVNTTIGQVSALQT 20
| | | | | | | | | | | | | | | | | | | | | |
Db 48 QQAQAQVNTTIGQVSALQT 67

RESULT 4
US-10-383-930-30
; Sequence 30, Application US/10383930
; Publication No. US20040127400A1
; GENERAL INFORMATION:
; APPLICANT: Smith, Daniel J
; APPLICANT: Taubman, Martin A
; TITLE OF INVENTION: Immunogenicity of Glucan Binding Protein
; FILE REFERENCE: 25669-018
; CURRENT APPLICATION NUMBER: US/10/383,930
; CURRENT FILING DATE: 2003-03-07
; PRIOR APPLICATION NUMBER: 60/402,483
; PRIOR FILING DATE: 2002-08-08
; PRIOR APPLICATION NUMBER: 60/363,209
; PRIOR FILING DATE: 2002-03-07
; NUMBER OF SEQ ID NOS: 41
; SOFTWARE: PatentIn version 3.2

; SEQ ID NO 30
; LENGTH: 431
; TYPE: PRT
; ORGANISM: Streptococcus mutans
US-10-383-930-30

Query Match 100.0%; Score 93; DB 4; Length 431;
Best Local Similarity 100.0%; Pred. No. 2.3e-06;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 QQAQAQVNTTIGQVSALQT 20
| | | | | | | | | | | | | | | | | | | | | |
Db 48 QQAQAQVNTTIGQVSALQT 67

RESULT 5
US-10-383-930-33
; Sequence 33, Application US/10383930
; Publication No. US20040127400A1
; GENERAL INFORMATION:
; APPLICANT: Smith, Daniel J
; APPLICANT: Taubman, Martin A
; TITLE OF INVENTION: Immunogenicity of Glucan Binding Protein
; FILE REFERENCE: 25669-018
; CURRENT APPLICATION NUMBER: US/10/383,930
; CURRENT FILING DATE: 2003-03-07
; PRIOR APPLICATION NUMBER: 60/402,483
; PRIOR FILING DATE: 2002-08-08
; PRIOR APPLICATION NUMBER: 60/363,209
; PRIOR FILING DATE: 2002-03-07
; NUMBER OF SEQ ID NOS: 41
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 33
; LENGTH: 431
; TYPE: PRT
; ORGANISM: Streptococcus mutans
US-10-383-930-33

Query Match 100.0%; Score 93; DB 4; Length 431;
Best Local Similarity 100.0%; Pred. No. 2.3e-06;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 QQAQAQVNTTIGQVSALQT 20
| | | | | | | | | | | | | | | | | | | | | |
Db 48 QQAQAQVNTTIGQVSALQT 67

RESULT 6
US-10-797-821-29
; Sequence 29, Application US/10797821
; Publication No. US20050031633A1
; GENERAL INFORMATION:
; APPLICANT: Smith, Daniel J.
; APPLICANT: Taubman, Martin A.
; TITLE OF INVENTION: Glucan Binding Protein and Glycosyltransferase Immunogens
; FILE REFERENCE: 25669-020
; CURRENT APPLICATION NUMBER: US/10/797,821
; CURRENT FILING DATE: 2004-03-09
; PRIOR APPLICATION NUMBER: 10/383,930
; PRIOR FILING DATE: 2003-03-07
; PRIOR APPLICATION NUMBER: 60/363,209
; PRIOR FILING DATE: 2002-03-07
; PRIOR APPLICATION NUMBER: 60/402,483
; PRIOR FILING DATE: 2002-08-08
; PRIOR APPLICATION NUMBER: 09/290,049
; PRIOR FILING DATE: 1999-04-12
; PRIOR APPLICATION NUMBER: 60/081,550
; PRIOR FILING DATE: 1998-04-13
; PRIOR APPLICATION NUMBER: 60/115,142
; PRIOR FILING DATE: 1999-01-08
; NUMBER OF SEQ ID NOS: 45
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 29

```
; LENGTH: 431
; TYPE: PRT
; ORGANISM: Streptococcus mutans
US-10-797-821-29

Query Match      100.0%; Score 93; DB 5; Length 431;
Best Local Similarity 100.0%; Pred. No. 2.3e-06;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 QQAQAQVNTTIGQVSALQT 20
   |||||
Db 48 QQAQAQVNTTIGQVSALQT 67

RESULT 7
US-10-797-821-30
; Sequence 30, Application US/10797821
; Publication No. US20050031633A1
; GENERAL INFORMATION:
; APPLICANT: Smith, Daniel J.
; APPLICANT: Taubman, Martin A.
; TITLE OF INVENTION: Glucan Binding Protein and Glycosyltransferase Immunogens
; CURRENT APPLICATION NUMBER: US/10/797,821
; CURRENT FILING DATE: 2004-03-09
; PRIOR APPLICATION NUMBER: 10/383,930
; PRIOR FILING DATE: 2003-03-07
; PRIOR APPLICATION NUMBER: 60/363,209
; PRIOR FILING DATE: 2002-03-07
; PRIOR APPLICATION NUMBER: 60/402,483
; PRIOR FILING DATE: 2002-08-08
; PRIOR APPLICATION NUMBER: 09/290,049
; PRIOR FILING DATE: 1999-04-12
; PRIOR APPLICATION NUMBER: 60/081,550
; PRIOR FILING DATE: 1998-04-13
; PRIOR APPLICATION NUMBER: 60/115,142
; PRIOR FILING DATE: 1999-01-08
; NUMBER OF SEQ ID NOS: 45
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 30
; LENGTH: 431
; TYPE: PRT
; ORGANISM: Streptococcus mutans
US-10-797-821-30

Query Match      100.0%; Score 93; DB 5; Length 431;
Best Local Similarity 100.0%; Pred. No. 2.3e-06;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 QQAQAQVNTTIGQVSALQT 20
   |||||
Db 48 QQAQAQVNTTIGQVSALQT 67

RESULT 8
US-10-797-821-33
; Sequence 33, Application US/10797821
; Publication No. US20050031633A1
; GENERAL INFORMATION:
; APPLICANT: Smith, Daniel J.
; APPLICANT: Taubman, Martin A.
; TITLE OF INVENTION: Glucan Binding Protein and Glycosyltransferase Immunogens
; CURRENT APPLICATION NUMBER: US/10/797,821
; CURRENT FILING DATE: 2004-03-09
; PRIOR APPLICATION NUMBER: 10/383,930
; PRIOR FILING DATE: 2003-03-07
; PRIOR APPLICATION NUMBER: 60/363,209
; PRIOR FILING DATE: 2002-03-07
; PRIOR APPLICATION NUMBER: 60/402,483
; PRIOR FILING DATE: 2002-08-08
; PRIOR APPLICATION NUMBER: 09/290,049
; PRIOR FILING DATE: 1999-04-12
```

```
; PRIOR APPLICATION NUMBER: 60/081,550
; PRIOR FILING DATE: 1998-04-13
; PRIOR APPLICATION NUMBER: 60/115,142
; PRIOR FILING DATE: 1999-01-08
; NUMBER OF SEQ ID NOS: 45
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 33
; LENGTH: 431
; TYPE: PRT
; ORGANISM: Streptococcus mutans
US-10-797-821-33

Query Match      100.0%; Score 93; DB 5; Length 431;
Best Local Similarity 100.0%; Pred. No. 2.3e-06;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 QQAQAQVNTTIGQVSALQT 20
   |||||
Db 48 QQAQAQVNTTIGQVSALQT 67

RESULT 9
US-10-383-930-31
; Sequence 31, Application US/10383930
; Publication No. US20040127400A1
; GENERAL INFORMATION:
; APPLICANT: Smith, Daniel J
; APPLICANT: Taubman, Martin A
; TITLE OF INVENTION: Immunogenicity of Glucan Binding Protein
; FILE REFERENCE: 25669-018
; CURRENT APPLICATION NUMBER: US/10/383,930
; CURRENT FILING DATE: 2003-03-07
; PRIOR APPLICATION NUMBER: 60/402,483
; PRIOR FILING DATE: 2002-08-08
; PRIOR APPLICATION NUMBER: 60/363,209
; PRIOR FILING DATE: 2002-03-07
; NUMBER OF SEQ ID NOS: 41
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 31
; LENGTH: 432
; TYPE: PRT
; ORGANISM: Streptococcus mutans
US-10-383-930-31

Query Match      100.0%; Score 93; DB 4; Length 432;
Best Local Similarity 100.0%; Pred. No. 2.3e-06;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 QQAQAQVNTTIGQVSALQT 20
   |||||
Db 48 QQAQAQVNTTIGQVSALQT 67

RESULT 10
US-10-383-930-32
; Sequence 32, Application US/10383930
; Publication No. US20040127400A1
; GENERAL INFORMATION:
; APPLICANT: Smith, Daniel J
; APPLICANT: Taubman, Martin A
; TITLE OF INVENTION: Immunogenicity of Glucan Binding Protein
; FILE REFERENCE: 25669-018
; CURRENT APPLICATION NUMBER: US/10/383,930
; CURRENT FILING DATE: 2003-03-07
; PRIOR APPLICATION NUMBER: 60/402,483
; PRIOR FILING DATE: 2002-08-08
; PRIOR APPLICATION NUMBER: 60/363,209
; PRIOR FILING DATE: 2002-03-07
; NUMBER OF SEQ ID NOS: 41
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 32
; LENGTH: 432
; TYPE: PRT
```

```
; ORGANISM: Streptococcus mutans
US-10-383-930-32

Query Match      100.0%; Score 93; DB 4; Length 432;
Best Local Similarity 100.0%; Pred. No. 2.3e-06;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 QQAQAQVNTTIGQVSALQT 20
Db 48 QQAQAQVNTTIGQVSALQT 67

RESULT 11
US-10-797-821-31
; Sequence 31, Application US/10797821
; Publication No. US20050031633A1
; GENERAL INFORMATION:
; APPLICANT: Smith, Daniel J.
; APPLICANT: Taubman, Martin A.
; TITLE OF INVENTION: Glucan Binding Protein and Glycosyltransferase Immunogens
; FILE REFERENCE: 25669-020
; CURRENT APPLICATION NUMBER: US/10/797,821
; CURRENT FILING DATE: 2004-03-09
; PRIOR APPLICATION NUMBER: 10/383,930
; PRIOR FILING DATE: 2003-03-07
; PRIOR APPLICATION NUMBER: 60/363,209
; PRIOR FILING DATE: 2002-03-07
; PRIOR APPLICATION NUMBER: 60/402,483
; PRIOR FILING DATE: 2002-08-08
; PRIOR APPLICATION NUMBER: 09/290,049
; PRIOR FILING DATE: 1999-04-12
; PRIOR APPLICATION NUMBER: 60/081,550
; PRIOR FILING DATE: 1998-04-13
; PRIOR APPLICATION NUMBER: 60/115,142
; NUMBER OF SEQ ID NOS: 45
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 31
; LENGTH: 432
; TYPE: PRT
; ORGANISM: Streptococcus mutans
US-10-797-821-31

Query Match      100.0%; Score 93; DB 5; Length 432;
Best Local Similarity 100.0%; Pred. No. 2.3e-06;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 QQAQAQVNTTIGQVSALQT 20
Db 48 QQAQAQVNTTIGQVSALQT 67

RESULT 12
US-10-797-821-32
; Sequence 32, Application US/10797821
; Publication No. US20050031633A1
; GENERAL INFORMATION:
; APPLICANT: Smith, Daniel J.
; APPLICANT: Taubman, Martin A.
; TITLE OF INVENTION: Glucan Binding Protein and Glycosyltransferase Immunogens
; FILE REFERENCE: 25669-020
; CURRENT APPLICATION NUMBER: US/10/797,821
; CURRENT FILING DATE: 2004-03-09
; PRIOR APPLICATION NUMBER: 10/383,930
; PRIOR FILING DATE: 2003-03-07
; PRIOR APPLICATION NUMBER: 60/363,209
; PRIOR FILING DATE: 2002-03-07
; PRIOR APPLICATION NUMBER: 60/402,483
; PRIOR FILING DATE: 2002-08-08
; PRIOR APPLICATION NUMBER: 09/290,049
; PRIOR FILING DATE: 1999-04-12
; PRIOR APPLICATION NUMBER: 60/081,550
; PRIOR FILING DATE: 1998-04-13
```

```
; PRIOR APPLICATION NUMBER: 60/115,142
; PRIOR FILING DATE: 1999-01-08
; NUMBER OF SEQ ID NOS: 45
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 32
; LENGTH: 432
; TYPE: PRT
; ORGANISM: Streptococcus mutans
US-10-797-821-32

Query Match      100.0%; Score 93; DB 5; Length 432;
Best Local Similarity 100.0%; Pred. No. 2.3e-06;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 QQAQAQVNTTIGQVSALQT 20
Db 48 QQAQAQVNTTIGQVSALQT 67

RESULT 13
US-10-383-930-9
; Sequence 9, Application US/10383930
; Publication No. US20040127400A1
; GENERAL INFORMATION:
; APPLICANT: Smith, Daniel J.
; APPLICANT: Taubman, Martin A.
; TITLE OF INVENTION: Immunogenicity of Glucan Binding Protein
; FILE REFERENCE: 25669-018
; CURRENT APPLICATION NUMBER: US/10/383,930
; CURRENT FILING DATE: 2003-03-07
; PRIOR APPLICATION NUMBER: 60/402,483
; PRIOR FILING DATE: 2002-08-08
; PRIOR APPLICATION NUMBER: 60/363,209
; PRIOR FILING DATE: 2002-03-07
; NUMBER OF SEQ ID NOS: 41
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 9
; LENGTH: 20
; TYPE: PRT
; ORGANISM: Streptococcus mutans
US-10-383-930-9

Query Match      80.6%; Score 75; DB 4; Length 20;
Best Local Similarity 100.0%; Pred. No. 5.6e-05;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 5 QAQVNTTIGQVSALQT 20
Db 1 QAQVNTTIGQVSALQT 16

RESULT 14
US-10-797-821-9
; Sequence 9, Application US/10797821
; Publication No. US20050031633A1
; GENERAL INFORMATION:
; APPLICANT: Smith, Daniel J.
; APPLICANT: Taubman, Martin A.
; TITLE OF INVENTION: Glucan Binding Protein and Glycosyltransferase Immunogens
; FILE REFERENCE: 25669-020
; CURRENT APPLICATION NUMBER: US/10/797,821
; CURRENT FILING DATE: 2004-03-09
; PRIOR APPLICATION NUMBER: 10/383,930
; PRIOR FILING DATE: 2003-03-07
; PRIOR APPLICATION NUMBER: 60/363,209
; PRIOR FILING DATE: 2002-03-07
; PRIOR APPLICATION NUMBER: 60/402,483
; PRIOR FILING DATE: 2002-08-08
; PRIOR APPLICATION NUMBER: 09/290,049
; PRIOR FILING DATE: 1999-04-12
; PRIOR APPLICATION NUMBER: 60/081,550
; PRIOR FILING DATE: 1998-04-13
; PRIOR APPLICATION NUMBER: 60/115,142
```

; PRIOR FILING DATE: 1999-01-08
 ; NUMBER OF SEQ ID NOS: 45
 ; SOFTWARE: PatentIn version 3.2
 ; SEQ ID NO 9
 ; LENGTH: 20
 ; TYPE: PRT
 ; ORGANISM: Artificial
 ; FEATURE:
 ; OTHER INFORMATION: GbpB peptide
 US-10-797-821-9

Query Match 80.6%; Score 75; DB 5; Length 20;
 Best Local Similarity 100.0%; Pred. No. 5.6e-05;
 Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 5 QAQVNTIQGVSAALQT 20
 |||||
 Db 1 QAQVNTIQGVSAALQT 16

RESULT 15
 US-10-474-792-600
 ; Sequence 600, Application US/10474792
 ; Publication No. US20040236072A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Olmsted, Stephen
 ; APPLICANT: Zagursky, Robert
 ; APPLICANT: Nickbarg, Elliot
 ; APPLICANT: Winter, Lourie
 ; TITLE OF INVENTION: SURFACE PROTEINS OF STREPTOCOCCUS PYOGENES
 ; FILE REFERENCE: AM 100399
 ; CURRENT APPLICATION NUMBER: US/10/474,792
 ; CURRENT FILING DATE: 2003-10-14
 ; NUMBER OF SEQ ID NOS: 674
 ; SOFTWARE: PatentIn version 3.0
 ; SEQ ID NO 600
 ; LENGTH: 398
 ; TYPE: PRT
 ; ORGANISM: Streptococcus pyogenes
 US-10-474-792-600

Query Match 62.4%; Score 58; DB 5; Length 398;
 Best Local Similarity 60.0%; Pred. No. 0.79;
 Matches 12; Conservative 5; Mismatches 3; Indels 0; Gaps 0;

QY 1 QQAQAQVNTIQGVSAALQT 20
 |||||
 Db 45 QKAQNVSAALQAQVSSLSQS 64

Search completed: January 25, 2006, 20:10:04
 Job time : 61 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2006 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: January 25, 2006, 19:15:26 ; Search time 5.5 Seconds
(without alignments)
39.378 Million cell updates/sec

Title: US-10-797-821-8
Perfect score: 93
Sequence: 1 QQAQAQVNTIQGVSAALQT 20

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 75621 seqs, 10829074 residues

Total number of hits satisfying chosen parameters: 75621

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published Applications AA New:
1: /cgn2_6/ptodata/2/pubpaa/US08_NEW_PUB.pap:
2: /cgn2_6/ptodata/2/pubpaa/US06_NEW_PUB.pap:
3: /cgn2_6/ptodata/2/pubpaa/US07_NEW_PUB.pap:
4: /cgn2_6/ptodata/2/pubpaa/ECT_NEW_PUB.pap:
5: /cgn2_6/ptodata/2/pubpaa/US09_NEW_PUB.pap:
6: /cgn2_6/ptodata/2/pubpaa/US10_NEW_PUB.pap:
7: /cgn2_6/ptodata/2/pubpaa/US11_NEW_PUB.pap:
8: /cgn2_6/ptodata/2/pubpaa/US60_NEW_PUB.pap:

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	93	100.0	431	7 US-11-052-554A-210	Sequence 210, App
2	58	62.4	398	7 US-11-052-554A-252	Sequence 252, App
3	44	47.3	472	6 US-10-508-263-62	Sequence 62, Appl
4	44	47.3	472	6 US-10-508-263-64	Sequence 64, Appl
5	44	47.3	488	6 US-10-508-263-12	Sequence 12, Appl
6	44	47.3	490	6 US-10-508-263-10	Sequence 10, Appl
7	42	45.2	496	6 US-10-508-263-14	Sequence 14, Appl
8	42	45.2	515	7 US-11-132-142-8	Sequence 8, Appl
9	41	44.1	1122	6 US-10-467-657-6112	Sequence 6112, Ap
10	40	43.0	1995	7 US-11-069-834-60	Sequence 60, Appl
11	39	41.9	455	6 US-10-508-263-66	Sequence 66, Appl
12	39	41.9	509	6 US-10-508-263-18	Sequence 18, Appl
13	39	41.9	544	6 US-10-821-234-889	Sequence 889, App
14	39	41.9	2101	6 US-10-857-780-23	Sequence 23, Appl
15	39	41.9	3674	7 US-11-000-463-454	Sequence 454, App
16	38	40.9	175	7 US-11-113-581-88	Sequence 88, Appl
17	38	40.9	175	7 US-11-113-581-97	Sequence 97, Appl
18	38	40.9	184	6 US-10-508-263-16	Sequence 16, Appl
19	38	40.9	210	7 US-11-052-554A-328	Sequence 328, App
20	38	40.9	264	7 US-11-188-743-24	Sequence 24, Appl
21	38	40.9	264	7 US-11-188-743-25	Sequence 25, Appl
22	38	40.9	364	7 US-11-129-143-61	Sequence 61, Appl
23	38	40.9	476	6 US-10-821-234-1064	Sequence 1064, Ap
24	37	39.8	76	6 US-10-510-880-6	Sequence 6, Appl
25	37	39.8	76	6 US-10-977-334-2	Sequence 2, Appl

26	37	39.8	108	6 US-10-510-880-5	Sequence 5, Appli
27	37	39.8	108	6 US-10-875-800-1	Sequence 1, Appli
28	37	39.8	108	6 US-10-299-977-1	Sequence 1, Appli
29	37	39.8	134	6 US-10-510-880-1	Sequence 1, Appli
30	37	39.8	134	6 US-10-875-800-2	Sequence 2, Appli
31	37	39.8	134	7 US-11-043-590-12	Sequence 12, Appl
32	37	39.8	162	7 US-11-043-590-13	Sequence 13, Appl
33	37	39.8	165	5 US-09-978-360A-497	Sequence 497, App
34	37	39.8	302	7 US-11-156-084-334	Sequence 334, App
35	37	39.8	354	6 US-10-793-626-3098	Sequence 3098, Ap
36	37	39.8	451	6 US-10-508-263-68	Sequence 68, Appl
37	37	39.8	451	6 US-10-508-263-70	Sequence 70, Appl
38	37	39.8	498	6 US-10-510-947-9	Sequence 9, Appli
39	37	39.8	651	7 US-11-172-145-10	Sequence 10, Appl
40	37	39.8	834	7 US-11-052-554A-212	Sequence 212, App
41	37	39.8	1448	6 US-10-485-517-212	Sequence 212, App
42	37	39.8	1992	7 US-11-069-834-58	Sequence 58, Appl
43	37	39.8	2000	7 US-11-069-834-56	Sequence 56, Appl
44	36	38.7	116	7 US-11-008-570-122	Sequence 122, App
45	36	38.7	163	6 US-10-510-386-126	Sequence 126, App

ALIGNMENTS

RESULT 1
US-11-052-554A-210
; Sequence 210, Application US/11052554A
; Publication No. US20050288866A1
; GENERAL INFORMATION:
; APPLICANT: Sachdeva, et al.
; TITLE OF INVENTION: COMPUTATIONAL METHOD FOR IDENTIFYING ADHESIN AND ADHESIN-LIKE
; FILE REFERENCE: 30853/40359A
; CURRENT APPLICATION NUMBER: US/11/052.554A
; CURRENT FILING DATE: 2005-02-07
; PRIOR APPLICATION NUMBER: US 60/589,227
; PRIOR FILING DATE: 2004-07-20
; PRIOR APPLICATION NUMBER: IN 173/DEL/2004
; PRIOR FILING DATE: 2004-02-06
; NUMBER OF SEQ ID NOS: 763
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 210
; LENGTH: 431
; TYPE: PRT
; ORGANISM: Streptococcus mutans UA159
US-11-052-554A-210

Query Match 100.0%; Score 93; DB 7; Length 431;
Best Local Similarity 100.0%; Pred. No. 3.8e-08;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 QQAQAQVNTIQGVSAALQT 20
DB 48 QQAQAQVNTIQGVSAALQT 67

RESULT 2
US-11-052-554A-252
; Sequence 252, Application US/11052554A
; Publication No. US20050288866A1
; GENERAL INFORMATION:
; APPLICANT: Sachdeva, et al.
; TITLE OF INVENTION: COMPUTATIONAL METHOD FOR IDENTIFYING ADHESIN AND ADHESIN-LIKE
; FILE REFERENCE: 30853/40359A
; CURRENT APPLICATION NUMBER: US/11/052.554A
; CURRENT FILING DATE: 2005-02-07
; PRIOR APPLICATION NUMBER: US 60/589,227
; PRIOR FILING DATE: 2004-07-20
; PRIOR APPLICATION NUMBER: IN 173/DEL/2004
; PRIOR FILING DATE: 2004-02-06
; NUMBER OF SEQ ID NOS: 763

```

; Publication No. US20050260754A1
; GENERAL INFORMATION:
; APPLICANT: BASF Plant Science GmbH
; TITLE OF INVENTION: Constructs and methods for regulating gene expression
; FILE REFERENCE: 53262-20085.00
; CURRENT APPLICATION NUMBER: US/10/508,263
; CURRENT FILING DATE: 2004-09-20
; NUMBER OF SEQ ID NOS: 126
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 12
; LENGTH: 488
; TYPE: PRT
; ORGANISM: Brassica napus
US-10-508-263-12

Query Match          47.3%; Score 44; DB 6; Length 488;
Best Local Similarity 50.0%; Pred. No. 5;
Matches      8; Conservative      4; Mismatches      4; Indels      0; Gaps      0;

QY      4      AQAQVNTIQQVSALQ 19
         | | | | | : | : | :
DB      426      ANAQINTLAGRTSVLR 441

RESULT 6
US-10-508-263-10
; Sequence 10, Application US/10508263
; Publication No. US20050260754A1
; GENERAL INFORMATION:
; APPLICANT: BASF Plant Science GmbH
; TITLE OF INVENTION: Constructs and methods for regulating gene expression
; FILE REFERENCE: 53262-20085.00
; CURRENT APPLICATION NUMBER: US/10/508,263
; CURRENT FILING DATE: 2004-09-20
; NUMBER OF SEQ ID NOS: 126
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 10
; LENGTH: 490
; TYPE: PRT
; ORGANISM: Brassica napus
US-10-508-263-10

Query Match          47.3%; Score 44; DB 6; Length 490;
Best Local Similarity 50.0%; Pred. No. 5;
Matches      8; Conservative      4; Mismatches      4; Indels      0; Gaps      0;

QY      4      AQAQVNTIQQVSALQ 19
         | | | | | : | : | :
DB      428      ANAQINTLAGRTSVLR 443

RESULT 7
US-10-508-263-14
; Sequence 14, Application US/10508263
; Publication No. US20050260754A1
; GENERAL INFORMATION:
; APPLICANT: BASF Plant Science GmbH
; TITLE OF INVENTION: Constructs and methods for regulating gene expression
; FILE REFERENCE: 53262-20085.00
; CURRENT APPLICATION NUMBER: US/10/508,263
; CURRENT FILING DATE: 2004-09-20
; NUMBER OF SEQ ID NOS: 126
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 14
; LENGTH: 496
; TYPE: PRT
; ORGANISM: Brassica napus
US-10-508-263-14

Query Match          45.2%; Score 42; DB 6; Length 496;
Best Local Similarity 43.8%; Pred. No. 11;
Matches      7; Conservative      5; Mismatches      4; Indels      0; Gaps      0;

```


QY 4 AQAQVNTIQGVSAIQ 19
Db 434 ANAQINTLAGRTSVNR 449

RESULT 8
US-11-132-142-8
; Sequence 8, Application US/111132142
; Publication No. US20050276818A1
; GENERAL INFORMATION:
; APPLICANT: The Burnham Institute
; APPLICANT: Sikora, Sergey
; APPLICANT: Godzik, Adam
; TITLE OF INVENTION: UNCHARACTERIZED ORF3 IN SARS-CORONAVIRUS IS A
; FILE OF INVENTION: CYCLIC-AMP-DEPENDENT KINASE AND A TARGET FOR SARS THERAPY
; CURRENT APPLICATION NUMBER: US/11/132,142
; PRIOR FILING DATE: 2005-05-17
; PRIOR APPLICATION NUMBER: 60/571,698
; PRIOR FILING DATE: 2004-05-17
; NUMBER OF SEQ ID NOS: 21
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 8
; LENGTH: 515
; TYPE: PRT
; ORGANISM: Cryptococcus sp.
US-11-132-142-8

Query Match 45.2%; Score 42; DB 7; Length 515;
Best Local Similarity 50.0%; Pred. No. 11;
Matches 11; Conservative 4; Mismatches 3; Indels 4; Gaps 1;

QY 1 QQAQAQVNV---TIQGVSAI 18
Db 136 EQAQAQAQLHRVTQQAARVAAI 157

RESULT 9
US-10-467-657-6112
; Sequence 6112, Application US/10467657
; Publication No. US20050260581A1
; GENERAL INFORMATION:
; APPLICANT: CHIRON SPA
; APPLICANT: FONTANA Maria Rita
; APPLICANT: PIZZA Mariagrazia
; APPLICANT: MASIGNANI Vega
; APPLICANT: MONACI Elisabetta
; TITLE OF INVENTION: GONOCOCCAL PROTEINS AND NUCLEIC ACIDS
; FILE REFERENCE:
; CURRENT APPLICATION NUMBER: US/10/467,657
; CURRENT FILING DATE: 2003-08-11
; PRIOR APPLICATION NUMBER: GB-0103424.8
; PRIOR FILING DATE: 2001-02-12
; NUMBER OF SEQ ID NOS: 9218
; SOFTWARE: SeqWin99, version 1.04
; SEQ ID NO 6112
; LENGTH: 1122
; TYPE: PRT
; ORGANISM: Neisseria gonorrhoeae
US-10-467-657-6112

Query Match 44.1%; Score 41; DB 6; Length 1122;
Best Local Similarity 45.0%; Pred. No. 41;
Matches 9; Conservative 2; Mismatches 9; Indels 0; Gaps 0;

QY 1 QQAQAQVNTIQGVSAIQ 20
Db 219 QDETAAQVQALNEVHALQT 238

RESULT 10
US-11-069-834-60
; Sequence 60, Application US/11069834

; Publication No. US20050276811A1
; GENERAL INFORMATION:
; APPLICANT: CARROLL, MICHAEL C.
; APPLICANT: MOORE JR., FRANCIS D.
; APPLICANT: HECHTMAN, HERBERT B.
; TITLE OF INVENTION: NATURAL IGM ANTIBODIES AND INHIBITORS THEREOF
; FILE REFERENCE: CRA-002.01
; CURRENT APPLICATION NUMBER: US/11/069,834
; CURRENT FILING DATE: 2005-03-01
; PRIOR APPLICATION NUMBER: 60/588,648
; PRIOR FILING DATE: 2004-07-16
; PRIOR APPLICATION NUMBER: 60/549,123
; PRIOR FILING DATE: 2004-03-01
; NUMBER OF SEQ ID NOS: 65
; SOFTWARE: PatentIn Ver. 3.3
; SEQ ID NO 60
; LENGTH: 1995
; TYPE: PRT
; ORGANISM: Homo sapiens
US-11-069-834-60

Query Match 43.0%; Score 40; DB 7; Length 1995;
Best Local Similarity 38.9%; Pred. No. 1.2e+02;
Matches 7; Conservative 8; Mismatches 3; Indels 0; Gaps 0;

QY 3 AAQAQVNTIQGVSAIQ 20
Db 1245 ALEAEVSELRAELSLQ 1262

RESULT 11
US-10-508-263-66
; Sequence 66, Application US/10508263
; Publication No. US20050260754A1
; GENERAL INFORMATION:
; APPLICANT: BASF Plant Science GmbH
; TITLE OF INVENTION: Constructs and methods for regulating gene expression
; FILE REFERENCE: 53262-20085.00
; CURRENT APPLICATION NUMBER: US/10/508,263
; CURRENT FILING DATE: 2004-09-20
; NUMBER OF SEQ ID NOS: 126
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 66
; LENGTH: 455
; TYPE: PRT
; ORGANISM: Arabidopsis thaliana
US-10-508-263-66

Query Match 41.9%; Score 39; DB 6; Length 455;
Best Local Similarity 50.0%; Pred. No. 31;
Matches 7; Conservative 4; Mismatches 3; Indels 0; Gaps 0;

QY 6 AQVNTIQGVSAIQ 19
Db 399 AQVNTLAGRTSVNR 412

RESULT 12
US-10-508-263-18
; Sequence 18, Application US/10508263
; Publication No. US20050260754A1
; GENERAL INFORMATION:
; APPLICANT: BASF Plant Science GmbH
; TITLE OF INVENTION: Constructs and methods for regulating gene expression
; FILE REFERENCE: 53262-20085.00
; CURRENT APPLICATION NUMBER: US/10/508,263
; CURRENT FILING DATE: 2004-09-20
; NUMBER OF SEQ ID NOS: 126
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 18
; LENGTH: 509
; TYPE: PRT
; ORGANISM: Brassica napus

```

US-10-508-263-18
Query Match      41.9%; Score 39; DB 6; Length 509;
Best Local Similarity 50.0%; Pred. No. 35;
Matches 8; Conservative 4; Mismatches 4; Indels 0; Gaps 0;

QY 4 AQAQVNTIQGVSAQ 19
Db 447 ANAMVSTLAGRTSALR 462

RESULT 13
US-10-821-234-889
; Sequence 889, Application US/10821234
; Publication No. US20050255114A1
; GENERAL INFORMATION:
; APPLICANT: Labat, Ivan
; APPLICANT: Stache-Crain, Birgit
; APPLICANT: Andarmani, Susan
; APPLICANT: Tang, Y. Tom
; TITLE OF INVENTION: Methods for Diagnosis and Treatment of Preeclampsia
; FILE REFERENCE: 821A
; CURRENT APPLICATION NUMBER: US/10/821,234
; PRIOR FILING DATE: 2004-04-07
; PRIOR APPLICATION NUMBER: US 60/462,047
; PRIOR FILING DATE: 2003-04-07
; NUMBER OF SEQ ID NOS: 1704
; SOFTWARE: pt_seq_genes Version 1.0
; SEQ ID NO 889
; LENGTH: 544
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-821-234-889

Query Match      41.9%; Score 39; DB 6; Length 544;
Best Local Similarity 57.1%; Pred. No. 38;
Matches 8; Conservative 4; Mismatches 2; Indels 0; Gaps 0;

QY 6 AQVNTIQGVSAQ 19
Db 417 AOLSEIQTQISALE 430

RESULT 14
US-10-857-780-23
; Sequence 23, Application US/10857780
; Publication No. US20050272043A1
; GENERAL INFORMATION:
; APPLICANT: ROTH, RICHARD B.
; APPLICANT: BRAUN, ANDREAS
; APPLICANT: KAMMERER, STEFAN M.
; APPLICANT: NELSON, MATTHEW ROBERTS
; APPLICANT: RENELAND, RIKARD HENRY
; APPLICANT: ROYAL-WRIGHTSON, CAROLYN R.
; TITLE OF INVENTION: METHODS FOR IDENTIFYING RISK OF BREAST CANCER AND TREATMENTS
; FILE REFERENCE: THEREOF
; FILE REFERENCE: SEQ-4069-CP
; CURRENT APPLICATION NUMBER: US/10/857,780
; CURRENT FILING DATE: 2004-05-28
; PRIOR APPLICATION NUMBER: 10/723,681
; PRIOR FILING DATE: 2003-11-25
; PRIOR APPLICATION NUMBER: 60/490,234
; PRIOR FILING DATE: 2003-07-24
; PRIOR APPLICATION NUMBER: 60/525,239
; PRIOR FILING DATE: 2003-11-25
; NUMBER OF SEQ ID NOS: 4962
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 23
; LENGTH: 2101
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-857-780-23

```

```

Query Match      41.9%; Score 39; DB 6; Length 2101;
Best Local Similarity 53.3%; Pred. No. 1.8e+02;
Matches 8; Conservative 3; Mismatches 4; Indels 0; Gaps 0;

QY 2 QAAQAVNTIQGVQS 16
Db 1573 QRFOQLNELQAQLS 1587

RESULT 15
US-11-000-463-454
; Sequence 454, Application US/11000463
; Publication No. US20050266423A1
; GENERAL INFORMATION:
; APPLICANT: Tang, Y Tom
; APPLICANT: Liu, Chenghua
; APPLICANT: Asundi, Vinod
; APPLICANT: Chen, Rui-hong
; APPLICANT: Qian, Xiaohong B.
; APPLICANT: Wang, Zhiwei
; APPLICANT: Wehrman, Tom
; APPLICANT: Zhang, Jie
; APPLICANT: Zhou, Ping
; APPLICANT: Cao, Yi-Cheng
; APPLICANT: Dmanac, Radoje T.
; TITLE OF INVENTION: Novel Nucleic Acids and Polypeptides
; FILE REFERENCE: 785CIP4CN
; CURRENT APPLICATION NUMBER: US/11/000,463
; CURRENT FILING DATE: 2004-11-29
; PRIOR APPLICATION NUMBER: 10/291,265
; PRIOR FILING DATE: 2002-11-08
; PRIOR APPLICATION NUMBER: PCT/US01/02623
; PRIOR FILING DATE: 2001-01-25
; PRIOR APPLICATION NUMBER: 09/922,279
; PRIOR FILING DATE: 2001-08-03
; PRIOR APPLICATION NUMBER: 09/491,404
; PRIOR FILING DATE: 2000-01-25
; PRIOR APPLICATION NUMBER: 09/617,746
; PRIOR FILING DATE: 2000-07-17
; PRIOR APPLICATION NUMBER: 09/631,451
; PRIOR FILING DATE: 2000-08-03
; PRIOR APPLICATION NUMBER: 09/633,870
; PRIOR FILING DATE: 2000-09-15
; NUMBER OF SEQ ID NOS: 944
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 454
; LENGTH: 3674
; TYPE: PRT
; ORGANISM: Homo sapiens
US-11-000-463-454

Query Match      41.9%; Score 39; DB 7; Length 3674;
Best Local Similarity 44.4%; Pred. No. 3.5e+02;
Matches 8; Conservative 3; Mismatches 7; Indels 0; Gaps 0;

QY 2 QAAQAVNTIQGVSAQ 19
Db 1071 QPLOGQVETLQGLLRQV 1088

Search completed: January 25, 2006, 20:11:08
Job time : 5.5 secs

```

GenCore version 5.1.6
Copyright (c) 1993 - 2006 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: January 25, 2006, 18:48:24 ; Search time 68.2 Seconds
(without alignments)
128.850 Million cell updates/sec

Title: US-10-797-821-9
Perfect score: 94
Sequence: 1 QAQVNTIQQVSALQTOQAE 20

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 2443163 seqs, 439378781 residues

Total number of hits satisfying chosen parameters: 2443163

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : A_Geneseq_21.*
1: geneseqp1980s.*
2: geneseqp1990s.*
3: geneseqp2000s.*
4: geneseqp2001s.*
5: geneseqp2002s.*
6: geneseqp2003as.*
7: geneseqp2003bs.*
8: geneseqp2004s.*
9: geneseqp2005s.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	94	100.0	20	7	ADD93629 Streptoco
2	94	100.0	20	9	ADx37252 Streptoco
3	94	100.0	431	7	ADD93649 Streptoco
4	94	100.0	431	7	ADD93650 Streptoco
5	94	100.0	431	7	ADD93653 Streptoco
6	94	100.0	431	9	ADx37272 Streptoco
7	94	100.0	431	9	ADx37273 Streptoco
8	94	100.0	431	9	ADx37276 Streptoco
9	94	100.0	431	9	AEb91500 Microbial
10	94	100.0	432	7	AD93651 Streptoco
11	94	100.0	432	7	ADD93652 Streptoco
12	94	100.0	432	9	ADx37274 Streptoco
13	94	100.0	432	9	ADx37275 Streptoco
14	75	79.8	20	7	ADD93628 Streptoco
15	75	79.8	20	9	ADx37251 Streptoco
16	65	69.1	20	7	ADD93622 Streptoco
17	65	69.1	20	9	ADx37245 Streptoco
18	59	62.8	447	5	ABP29684 Streptoco
19	59	62.8	447	8	ADu69524 S agalact
20	59	62.8	447	8	ADv88392 Streptoco
21	59	62.8	447	8	ADv81808 Streptoco
22	59	62.8	447	8	ADv79645 Streptoco
23	52	55.3	210	2	AAY22579 Bacterial
24	52	55.3	392	6	ABU02747 S. pneumo

25	52	55.3	392	8	ADK47859	Adk47859 Streptoco
26	52	55.3	392	8	ADT50227	Adt50227 S_pneumon
27	52	55.3	392	8	ADT50226	Adt50226 S_pneumon
28	52	55.3	392	8	ADT50165	Adt50165 S_pneumon
29	52	55.3	399	8	ADR94595	Adr94595 Novel S.
30	52	55.3	399	9	AEA58465	Aea58465 Streptoco
31	51	54.3	398	5	ABP25919	Abp25919 Streptoco
32	51	54.3	398	8	ADR83884	Adr83884 S. pyogen
33	51	54.3	398	9	AEb91542	AEb91542 Microbial
34	51	54.3	456	5	ABB55584	Abb55584 Lactococc
35	48	51.1	395	5	ABP25918	Abp25918 Streptoco
36	47	50.0	461	2	AAR14530	Aar14530 Usp45 pro
37	47	50.0	461	2	AAR14150	Aar14150 MSP encod
38	47	50.0	1024	5	ABE04822	Abb04822 LDL recep
39	46	48.9	722	7	ABO74065	ABO74065 Pseudomon
40	46	48.9	768	2	AAY13477	Aay13477 Clone 039
41	45	47.9	383	7	ABO80875	ABO80875 Pseudomon
42	45	47.9	642	6	ABP79387	Abp79387 N. gonorr
43	45	47.9	642	6	ABU37367	Abu37367 Protein e
44	45	47.9	1024	5	ABB04820	Abb04820 LDL recep
45	45	47.9	1211	6	ABU41350	Abu41350 Protein e

ALIGNMENTS

RESULT 1
ADD93629
ID ADD93629 standard; peptide; 20 AA.
AC ADD93629;
XX
XX 29-JAN-2004 (first entry)
XX Streptococcus mutans glucan binding protein-B peptide fragment.
XX Glucan binding protein-B; GbpB; vaccine; anticaries; epitope; immunogen.
XX Streptococcus mutans.
XX WO2003075845-A2.
XX 18-SEP-2003.
XX 07-MAR-2003; 2003WO-US006962.
XX 07-MAR-2002; 2002US-0363209P.
XX 08-AUG-2002; 2002US-0402483P.
XX (FORS-) FORSYTH INST.
XX Smith DJ, Taubman MA;
XX WPI; 2003-845091/78.
Composition useful as vaccines for dental caries comprises a fragment of a glucan binding protein-B binding to a major histocompatibility complex class II protein.
Claim 11; Page 10; 49pp; English.

The present sequence is that of a peptide comprising amino acid residues 52-71 of the glucan binding protein-B (GbpB) of Streptococcus mutans strain SK32 ADD93649. The peptide binds to a major histocompatibility complex (MHC) class II protein. It was identified as a potential B cell epitope using a matrix-based algorithm for epitope prediction, which was used to search the primary amino acid sequence of GbpB for known MHC class II binding motifs. The peptide can be used in immunogenic compositions and subunit vaccines for dental caries. These comprise MHC class II protein-binding GbpB peptides covalently linked with peptide subunits (preferably from the catalytic domain) of a glucosyltransferase. The compositions are used in a claimed method of eliciting production of an antibody in a mammal. Diepitopic or multiepitopic polypeptides can be

CC prepared synthetically or by recombinant DNA technology. Antibodies
 CC raised against MHC class II binding fragments of GbpB can be used in
 CC passive immunisation.

SQ Sequence 20 AA;

Query Match 100.0%; Score 94; DB 7; Length 20;
 Best Local Similarity 100.0%; Pred. No. 1.8e-08;
 Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 QAVNTIQGVSAALQTQQAE 20
 |||||
 DB 1 QAVNTIQGVSAALQTQQAE 20
 |||||

RESULT 2

ADX37252
 ID ADX37252 standard; peptide; 20 AA.

XX
 AC ADX37252;

DT 21-APR-2005 (first entry)

DE Streptococcus mutant glucan binding protein B peptide #9.

KW immunogenicity; immune stimulation; glucan binding protein-B;
 KW microparticle; major histocompatibility complex; tooth disease.

XX Streptococcus mutans.

OS US2005031633-A1.

PN 10-FEB-2005.

PF 09-MAR-2004; 2004US-00797821.

PR 13-APR-1998; 98US-0081550P.

PR 08-JAN-1999; 99US-0115142P.

PR 12-APR-1999; 99US-00290049.

PR 07-MAR-2002; 2002US-0363209P.

PR 08-AUG-2002; 2002US-0402483P.

PR 07-MAR-2003; 2003US-00383930.

XX (SMIT/) SMITH D J.

PA (TAUB/) TAUBMAN M A.

XX Smith DJ, Taubman MA;

PI WPI; 2005-151644/16.

DR Claim 4; SEQ ID NO 9; 73pp; English.

XX The invention relates to a composition comprising a fragment of a glucan
 CC binding protein-B (GbpB) and a biocompatible microparticle, where the
 CC fragment binds to a major histocompatibility complex (MHC) class II
 CC protein. The composition is useful for producing an antibody for
 CC immunizing mammals against dental caries. This sequence corresponds to a
 CC Streptococcus mutans GbpB-derived peptide of the invention.

XX Sequence 20 AA;

Query Match 100.0%; Score 94; DB 9; Length 20;
 Best Local Similarity 100.0%; Pred. No. 1.8e-08;
 Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 QAVNTIQGVSAALQTQQAE 20
 |||||
 DB 1 QAVNTIQGVSAALQTQQAE 20
 |||||

RESULT 3
 ADD93649
 ID ADD93649 standard; protein; 431 AA.
 XX
 AC ADD93649;
 XX
 DT 29-JAN-2004 (first entry)
 XX
 DE Streptococcus mutans glucan binding protein-B.
 KW Glucan binding protein-B; GbpB; vaccine; anticaries; epitope; immunogen.
 XX Streptococcus mutans.
 OS
 XX
 FH Key Location/Qualifiers
 FT Region 6..25
 FT /note= "HLA-binding peptide"
 FT Region 16..35
 FT /note= "HLA-binding peptide"
 FT Region 33..52
 FT /note= "HLA-binding peptide"
 FT Region 37..56
 FT /note= "HLA-binding peptide"
 FT Region 48..67
 FT /note= "HLA-binding peptide"
 FT Region 52..71
 FT /note= "HLA-binding peptide"
 FT Region 88..107
 FT /note= "HLA-binding peptide"
 FT Region 113..132
 FT /note= "HLA-binding peptide"
 FT Region 117..136
 FT /note= "HLA-binding peptide"
 FT Region 137..156
 FT /note= "HLA-binding peptide"
 FT Region 174..193
 FT /note= "HLA-binding peptide"
 FT Region 194..213
 FT /note= "HLA-binding peptide"
 FT Region 214..233
 FT /note= "HLA-binding peptide"
 FT Region 248..267
 FT /note= "HLA-binding peptide"
 FT Region 289..308
 FT /note= "HLA-binding peptide"
 FT Region 306..325
 FT /note= "HLA-binding peptide"
 FT Region 311..330
 FT /note= "HLA-binding peptide"
 FT Region 349..368
 FT /note= "HLA-binding peptide"
 FT Region 365..384
 FT /note= "HLA-binding peptide"
 FT Region 383..402
 FT /note= "HLA-binding peptide"
 FT Region 403..422
 FT /note= "HLA-binding peptide"
 XX WO2003075845-A2.
 XX 18-SEP-2003.
 XX 07-MAR-2003; 2003WO-US006962.
 XX 07-MAR-2002; 2002US-0363209P.
 XX 08-AUG-2002; 2002US-0402483P.
 XX (FORS-) FORSYTH INST.
 XX Smith DJ, Taubman MA;

PF 09-MAR-2004; 2004US-00797821.
 XX
 PR 13-APR-1998; 98US-0081550P.
 PR 08-JAN-1999; 99US-0115142P.
 PR 12-APR-1999; 99US-00290049.
 PR 08-MAR-2002; 2002US-0363209P.
 PR 08-AUG-2002; 2002US-0402483P.
 PR 07-MAR-2003; 2003US-00383930.
 XX
 PA (SMIT/) SMITH D J.
 PA (TAUBS/) TAUBMAN M A.
 XX
 XX Smith DJ, Taubman MA;
 XX WPI; 2005-151644/16.
 DR
 XX
 XX New composition comprising a fragment of a glucan binding protein-B
 PT (GbpB) that binds to MHC class II protein, and a biocompatible
 PT microparticle, useful for producing an antibody (claimed) for immunizing
 PT mammals against dental caries.
 XX
 XX Claim 3; SEQ ID NO 33; 73pp; English.
 XX
 XX The invention relates to a composition comprising a fragment of a glucan
 CC binding protein-B (GbpB) and a biocompatible microparticle, where the
 CC fragment binds to a major histocompatibility complex (MHC) class II
 CC protein. The composition is useful for producing an antibody for
 CC immunizing mammals against dental caries. This sequence corresponds to a
 CC Streptococcus mutans GbpB protein of the invention.
 XX
 XX Sequence 431 AA;
 SQ
 Query Match 100.0%; Score 94; DB 9; Length 431;
 Best Local Similarity 100.0%; Pred. No. 7e-07;
 Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 QAVNTIQGVSAALQTQQAE 20
 Db 52 QAVNTIQGVSAALQTQQAE 71
 RESULT 9
 AEB91500
 ID AEB91500 standard; protein; 431 AA.
 XX
 AC AEB91500;
 XX
 DT 20-OCT-2005 (first entry)
 XX
 DE Microbial pathogen adhesin protein sequence, SEQ ID NO:210.
 XX
 KW algorithm; adhesin; pharmaceutical; vaccine; drug screening;
 KW bordetella pertussis infection; antibacterial; pneumonia;
 KW antiinflammatory; respiratory-gen.; gastric ulcer; antitumor;
 KW gastrointestinal-gen.; urinary tract infection; antimicrobial; uropathic.
 XX
 OS Streptococcus mutans.
 XX
 XX WO2005076010-A2.
 PN
 XX
 PD 18-AUG-2005.
 XX
 XX 07-FEB-2005; 2005WO-IN000037.
 PF
 XX 06-FEB-2004; 2004IN-DE000173.
 PR
 PR 20-JUL-2004; 2004US-0589227P.
 XX
 XX (COUL) COUNCIL SCI & IND RES SOUTH AFRICA.
 PA
 XX Sachdeva G, Kumar K, Jain P, Brahmachari SK, Ramachandran S;
 PI WPI; 2005-597835/61.
 XX
 DR
 XX

PT Computational method for identifying adhesin and adhesin like molecules,
 PT comprises computing sequence-based attributes of protein sequences using
 PT neural network software and training an artificial neural network.
 XX
 PS Claim 16; SEQ ID NO 210; 402pp; English.
 XX
 XX The present invention relates to a computational method (M1) for
 CC identifying adhesin and adhesin-like proteins, by computing the sequence-
 CC based attributes of protein sequences using five attribute modules of a
 CC neural network software, training an artificial neural network (ANN) for
 CC each of the computed five attributes, and identifying the adhesin and
 CC adhesin-like proteins having probability of being an adhesin (Pad) as
 CC equal or greater than 0.51. Also claimed is a set of 274 annotated genes
 CC encoding adhesin and adhesin-like proteins, having 274 fully defined 162-
 CC 11151 base pair (SEQ ID NO: 385-658) sequences; a set of 105 hypothetical
 CC genes encoding adhesin and adhesin-like proteins, having 105 fully
 CC defined 306-15876 base pairs (SEQ ID NO: 659-763); a set of 279 annotated
 CC adhesin and adhesin-like proteins, having 279 fully defined 53-3716 base
 CC pair (SEQ ID NO: 1-279); a set of 105 hypothetical adhesin and adhesin-
 CC like proteins, having 105 fully defined 106-5291 base pair (SEQ ID NO:
 CC 280-384) sequences; and a fully connected multilayer feed forward ANN (I)
 CC based on (M1). (M1) is useful for identifying adhesin and adhesin-like
 CC proteins, of therapeutic potential, and identifying and short-listing
 CC to eliminate diseases caused by various pathogenic organisms. (M1) is
 CC useful for identifying putative adhesins that are important in drug
 CC discovery and preventing therapeutics for whooping cough, pneumonia,
 CC gastric ulcer and urinary tract infections. (M1) identifies adhesins from
 CC distantly related organisms, and from bacteria belonging to a wide
 CC phylogenetic spectrum. (M1) is capable of predicting adhesive nature of
 CC unique proteins. The present sequence is a microbial pathogen adhesin
 CC protein sequence.
 XX
 XX Sequence 431 AA;
 SQ
 Query Match 100.0%; Score 94; DB 9; Length 431;
 Best Local Similarity 100.0%; Pred. No. 7e-07;
 Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 QAVNTIQGVSAALQTQQAE 20
 Db 52 QAVNTIQGVSAALQTQQAE 71
 RESULT 10
 ADD93651
 ID ADD93651 standard; protein; 432 AA.
 XX
 AC ADD93651;
 XX
 DT 29-JAN-2004 (first entry)
 XX
 DE Streptococcus mutans glucan binding protein-B.
 XX
 KW Glucan binding protein-B; GbpB; vaccine; anticaries; epitope; immunogen.
 XX
 OS Streptococcus mutans.
 XX
 XX WO2003075845-A2.
 PN
 XX
 PD 18-SEP-2003.
 XX
 XX 07-MAR-2003; 2003WO-US006962.
 PF
 XX 07-MAR-2002; 2002US-0363209P.
 PR
 PR 08-AUG-2002; 2002US-0402483P.
 XX
 XX (FORS-) FORSYTH INST.
 PA
 XX Smith DJ, Taubman MA;
 PI WPI; 2003-845091/78.
 XX
 DR GENBANK; AY046412.
 DR

CC	covalently linked with peptide subunits of a glucosyltransferase. The					
CC	compositions are used in a claimed method of eliciting production of an					
CC	antibody in a mammal. Dieptopic or multieptopic polypeptides can be					
CC	prepared synthetically or by recombinant DNA technology. Antibodies					
CC	raised against MHC class II binding fragments of GbpB can be used in					
CC	passive immunisation.					
XX						
SQ	Sequence 432 AA;					
	Query Match	100.0%;	Score 94;	DB 7;	Length 432;	
	Best Local Similarity	100.0%;	Pred. No. 7e-07;	Mismatches	0;	Gaps 0;
	Matches	20;	Conservative	0;	Indels	0;
QY	1	QAQVNTIQGVSAALQTQQAE	20			
Dd	52	QAQVNTIQGVSAALQTQQAE	71			
RESULT 12						
ADX37274						
ID	ADX37274	standard; protein; 432 AA.				
XX	AC	ADX37274;				
XX	AC					
DT	21-APR-2005	(first entry)				
XX						
DE	Streptococcus mutant glucan binding protein B variant #3.					
XX						
KW	immunogenicity; immune stimulation; glucan binding protein-B;					
KW	microparticle; major histocompatibility complex; tooth disease.					
XX						
OS	Streptococcus mutans.					
XX						
PN	US2005031633-A1.					
XX						
PD	10-FEB-2005.					
XX						
PF	09-MAR-2004; 2004US-00797821.					
XX						
PR	13-APR-1998; 98US-0081550P.					
PR	08-JAN-1999; 99US-0115142P.					
PR	12-APR-1999; 99US-00290049.					
PR	07-MAR-2002; 2002US-0363209P.					
PR	08-AUG-2002; 2002US-0402483P.					
PR	07-MAR-2003; 2003US-00383930.					
XX						
PA	(SMIT/) SMITH D J.					
PA	(TAUB/) TAUEMAN M A.					
XX						
PI	Smith DJ, Taubman MA;					
XX						
DR	WPI; 2005-151644/16.					
XX						
PT	New composition comprising a fragment of a glucan binding protein-B					
PT	(GbpB) that binds to MHC class II protein, and a biocompatible					
PT	microparticle, useful for producing an antibody (claimed) for immunizing					
PT	mammals against dental caries.					
XX						
PS	Claim 3; SEQ ID NO 31; 73pp; English.					
XX						
CC	The invention relates to a composition comprising a fragment of a glucan					
CC	binding protein-B (GbpB) and a biocompatible microparticle, where the					
CC	fragment binds to a major histocompatibility complex (MHC) class II					
CC	protein. The composition is useful for producing an antibody for					
CC	immunizing mammals against dental caries. This sequence corresponds to a					
CC	Streptococcus mutans gbpB protein of the invention.					
XX						
SQ	Sequence 432 AA;					
	Query Match	100.0%;	Score 94;	DB 9;	Length 432;	
	Best Local Similarity	100.0%;	Pred. No. 7e-07;	Mismatches	0;	Gaps 0;
	Matches	20;	Conservative	0;	Indels	0;


```

QY 1 QAAVNTIQGVSAALQTQQA 20
Db 52 QAAVNTIQGVSAALQTQQA 71

RESULT 13
ADX37275
ID ADX37275 standard; protein; 432 AA.
XX
AC ADX37275;
XX
DT 21-APR-2005 (first entry)
XX
DE Streptococcus mutant glucan binding protein B variant #4.
XX
KW immunogenicity; immune stimulation; glucan binding protein-B;
KW microparticle; major histocompatibility complex; tooth disease.
XX
OS Streptococcus mutans.
XX
PN US2005031633-A1.
XX
PD 10-FEB-2005.
XX
PF 09-MAR-2004; 2004US-00797821.
XX
PR 13-APR-1998; 98US-0081550P.
PR 08-JAN-1999; 99US-0115142P.
PR 12-APR-1999; 99US-00290049.
PR 07-MAR-2002; 2002US-0363209P.
PR 08-AUG-2002; 2002US-0402483P.
PR 07-MAR-2003; 2003US-00383930.
XX
PA (SMIT/) SMITH D J.
PA (TAUB/) TAUBMAN M A.
XX
PI Smith DJ, Taubman MA;
XX
DR WPI; 2005-151644/16.
XX
PT New composition comprising a fragment of a glucan binding protein-B
PT (GbpB) that binds to MHC class II protein, and a biocompatible
PT microparticle, useful for producing an antibody (claimed) for immunizing
PT mammals against dental caries.
XX
PS Claim 3; SEQ ID NO 32; 73pp; English.
XX
CC The invention relates to a composition comprising a fragment of a glucan
CC binding protein-B (GbpB) and a biocompatible microparticle, where the
CC fragment binds to a major histocompatibility complex (MHC) class II
CC protein. The composition is useful for producing an antibody for
CC immunizing mammals against dental caries. This sequence corresponds to a
CC Streptococcus mutans GbpB protein of the invention.
XX
SQ Sequence 432 AA;

Query Match 100.0%; Score 94; DB 9; Length 432;
Best Local Similarity 100.0%; Pred. No. 7e-07;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 QAAVNTIQGVSAALQTQQA 20
Db 52 QAAVNTIQGVSAALQTQQA 71

RESULT 14
ADD93628
ID ADD93628 standard; peptide; 20 AA.
XX
AC ADD93628;
XX
DT 29-JAN-2004 (first entry)
XX
XX

```

```

DE Streptococcus mutans glucan binding protein-B peptide fragment.
XX
KW Glucan binding protein-B; GbpB; vaccine; anticaries; epitope; immunogen.
XX
OS Streptococcus mutans.
XX
PN WO2003075845-A2.
XX
PD 18-SEP-2003.
XX
PF 07-MAR-2003; 2003WO-US006962.
XX
PR 07-MAR-2002; 2002US-0363209P.
PR 08-AUG-2002; 2002US-0402483P.
XX
PA (FORS-) FORSYTH INST.
XX
PI Smith DJ, Taubman MA;
XX
DR WPI; 2003-845091/78.
XX
PT Composition useful as vaccines for dental caries comprises a fragment of
PT a glucan binding protein-B binding to a major histocompatibility complex
PT class II protein.
XX
PS Claim 11; Page 10; 49pp; English.
XX
CC The present sequence is that of a peptide comprising amino acid residues
CC 48-67 of the glucan binding protein-B (GbpB) of Streptococcus mutans
CC strain SK32 ADD93649. The peptide binds to a major histocompatibility
CC complex (MHC) class II protein. It was identified as a potential B cell
CC epitope using a matrix-based algorithm for epitope prediction, which was
CC used to search the primary amino acid sequence of GbpB for known MHC
CC class II binding motifs. The peptide can be used in immunogenic
CC compositions and subunit vaccines for dental caries. These comprise MHC
CC class II protein-binding GbpB peptides covalently linked with peptide
CC subunits (preferably from the catalytic domain) of a glucosyltransferase.
CC The compositions are used in a claimed method of eliciting production of
CC an antibody in a mammal. Dieptopic or multi-epitopic polypeptides can be
CC prepared synthetically or by recombinant DNA technology. Antibodies
CC raised against MHC class II binding fragments of GbpB can be used in
CC passive immunisation.
XX
SQ Sequence 20 AA;

Query Match 79.8%; Score 75; DB 7; Length 20;
Best Local Similarity 100.0%; Pred. No. 2.6e-05;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 QAAVNTIQGVSAALQT 16
Db 5 QAAVNTIQGVSAALQT 20

RESULT 15
ADX37251
ID ADX37251 standard; peptide; 20 AA.
XX
AC ADX37251;
XX
DT 21-APR-2005 (first entry)
XX
DE Streptococcus mutant glucan binding protein B peptide #8.
XX
KW immunogenicity; immune stimulation; glucan binding protein-B;
KW microparticle; major histocompatibility complex; tooth disease.
XX
OS Streptococcus mutans.
XX
PN US2005031633-A1.
XX
PD 10-FEB-2005.
XX

```

PF 09-MAR-2004; 2004US-00797821.
XX
PR 13-APR-1998; 98US-0081550P.
PR 08-JAN-1999; 99US-0115142P.
PR 12-APR-1999; 99US-00290049.
PR 07-MAR-2002; 2002US-0363209P.
PR 08-AUG-2002; 2002US-0402483P.
PR 07-MAR-2003; 2003US-00383930.
XX
XX (SMIT/) SMITH D J.
PA (TAUB/) TAUBMAN M A.
XX
PI Smith DJ, Taubman MA;
XX
DR WPI; 2005-151644/16.
XX
XX New composition comprising a fragment of a glucan binding protein-B
PT (GbpB) that binds to MHC class II protein, and a biocompatible
PT microparticle, useful for producing an antibody (claimed) for immunizing
PT mammals against dental caries.
XX
PS Claim 4; SEQ ID NO 8; 73pp; English.
XX
CC The invention relates to a composition comprising a fragment of a glucan
CC binding protein-B (GbpB) and a biocompatible microparticle, where the
CC fragment binds to a major histocompatibility complex (MHC) class II
CC protein. The composition is useful for producing an antibody for
CC immunizing mammals against dental caries. This sequence corresponds to a
CC Streptococcus mutans GbpB-derived peptide of the invention.
XX
SQ Sequence 20 AA;

Query Match 79.8%; Score 75; DB 9; Length 20;
Best Local Similarity 100.0%; Pred. NO. 2.6e-05;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 QAQVNTIQGVSAALQT 16
DB 5 QAQVNTIQGVSAALQT 20

Search completed: January 25, 2006, 19:01:45
Job time : 69.2 secs

Query Match 55.3%; Score 52; DB 2; Length 392;
Best Local Similarity 57.9%; Pred. No. 1.2;
Matches 11; Conservative 4; Mismatches 4; Indels 0; Gaps 0;

QY 1 QAQVNTIQGVSAALQTQQA 19
|||:|||||:|:
DB 52 QKQVDQIQGVSAIQAEQS 70
|||:|||||:|:

RESULT 3

B98124
general stress protein GSP-781 [imported] - Streptococcus pneumoniae (strain R6)
C:Species: Streptococcus pneumoniae
C>Date: 22-Oct-2001 #sequence_revision 22-Oct-2001 #text_change 09-Jul-2004
C:Accession: B98124
R:Hoskins, J.A.; Alborn Jr., W.; Arnold, J.; Blaszcak, L.; Burgett, S.; DeHoff, B.S.; E
e, R.; LeBlanc, D.J.; Lee, L.N.; Lefkowitz, E.J.; Lu, J.; Matsushima, P.; McAhren, S.; M
Y, P.; Sun, P.M.; Winkler, M.E.
J. Bacteriol. 183, 5709-5717, 2001
A:Authors: Yang, Y.; Young-Bellido, M.; Zhao, G.; Zook, C.; Baltz, R.H.; Jaskunas, S.R.;
A:Title: Genome of the Bacterium Streptococcus pneumoniae Strain R6.
A:Reference number: A37872; MUID:21429245; PMID:11544234
A:Accession: B98124
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-392 <KUR>
A:Cross-references: UNIPROT:Q97N55; UNIPROT:Q8DMY4; UNIPARC:UPI0000051B81; GB:AE007317;
C:Genetics:
A:Gene: gsp-781

Query Match 55.3%; Score 52; DB 2; Length 392;
Best Local Similarity 57.9%; Pred. No. 1.2;
Matches 11; Conservative 4; Mismatches 4; Indels 0; Gaps 0;

QY 1 QAQVNTIQGVSAALQTQQA 19
|||:|||||:|:
DB 52 QKQVDQIQGVSAIQAEQS 70
|||:|||||:|:

RESULT 4

B86903
hypothetical protein usp45 [imported] - Lactococcus lactis subsp. lactis (strain IL1403)
C:Species: Lactococcus lactis subsp. lactis
C>Date: 23-Mar-2001 #sequence_revision 23-Mar-2001 #text_change 09-Jul-2004
C:Accession: B86903
R:Botolin, A.; Wincker, P.; Mauger, S.; Jaillon, O.; Malarne, K.; Weissenbach, J.; Ehrlich
Genome Res. 11, 731-753, 2001
A:Title: The complete genome sequence of the lactic acid bacterium Lactococcus lactis se
A:Reference number: A86625; MUID:21235186; PMID:11337471
A:Accession: B86903
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-456 <STO>
A:Cross-references: UNIPROT:Q9CDJ1; UNIPARC:UPI0000006BFC; GB:AE0051176; PID:g12725296; E
A:Experimental source: strain IL1403
C:Genetics:
A:Gene: usp45

Query Match 54.3%; Score 51; DB 2; Length 456;
Best Local Similarity 52.6%; Pred. No. 2;
Matches 10; Conservative 6; Mismatches 3; Indels 0; Gaps 0;

QY 1 QAQVNTIQGVSAALQTQQA 19
|||:|||||:|:
DB 51 QAQVDSLQSKVDSLQKQA 69
|||:|||||:|:

RESULT 5

S05542
hypothetical protein, 54K - Enterococcus faecium
C:Species: Enterococcus faecium
C>Date: 07-Sep-1990 #sequence_revision 07-Sep-1990 #text_change 09-Jul-2004
C:Accession: S05542

R:Fuerst, P.; Moesch, H.U.; Solioz, M.
Nucleic Acids Res. 17, 6724, 1989

A:Title: A protein of unusual composition from Enterococcus faecium.
A:Reference number: S05542; MUID:8938598; PMID:2780297

A:Accession: S05542

A:Molecule type: DNA

A:Residues: 1-507 <FUE>

A:Cross-references: UNIPROT:P13692; UNIPARC:UPI000016FGFC; GB:X16421; EMBL:M26048; NID:9
A:Note: the authors translated the codon CGT for residues 221 and 223 as Lys

Query Match 52.1%; Score 49; DB 2; Length 507;
Best Local Similarity 47.1%; Pred. No. 4.6;
Matches 8; Conservative 6; Mismatches 3; Indels 0; Gaps 0;

QY 1 QAQVNTIQGVSAALQTQ 17
|||:|||||:|:
DB 43 QSQIEALQGVSAINTK 59
|||:|||||:|:

RESULT 6

JN0097
secreted 45K protein precursor - Lactococcus lactis
C:Species: Lactococcus lactis
C>Date: 16-Jul-1999 #sequence_revision 16-Jul-1999 #text_change 16-Jul-1999
C:Accession: JN0097
R:van Asseldonk, M.; Rutten, G.; Oteman, M.; Siezen, R.J.; de Vos, W.M.; Simons, G.
Gene 95, 155-160, 1990
A:Title: Cloning of usp45, a gene encoding a secreted protein from Lactococcus lactis su
A:Reference number: JN0097; MUID:91071599; PMID:2123812
A:Accession: JN0097
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-461 <VAN>
A:Cross-references: UNIPARC:UPI000016D72B; GB:M35374

Query Match 50.0%; Score 47; DB 2; Length 461;
Best Local Similarity 50.0%; Pred. No. 8.8;
Matches 9; Conservative 6; Mismatches 3; Indels 0; Gaps 0;

QY 1 QAQVNTIQGVSAALQTQ 18
|||:|||||:|:
DB 51 QAQVDSLQSKVDSLQKQ 68
|||:|||||:|:

RESULT 7

G91038
hypothetical protein ECS3279 [imported] - Escherichia coli (strain O157:H7, substrain RJ
C:Species: Escherichia coli
C>Date: 18-Jul-2001 #sequence_revision 18-Jul-2001 #text_change 09-Jul-2004
C:Accession: G91038
R:Hayashi, T.; Makino, K.; Ohnishi, M.; Kurokawa, K.; Ishii, K.; Yokoyama, K.; Han, C.G.
gasawara, N.; Yasunaga, T.; Kuhara, S.; Shiba, T.; Hatcori, M.; Shingawa, H.
DNA Res. 8, 11-22, 2001
A:Title: Complete genome sequence of enterohemorrhagic Escherichia coli O157:H7 and gene
A:Reference number: A99629; MUID:21156231; PMID:11258796
A:Accession: G91038
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-113 <HAY>
A:Cross-references: UNIPROT:Q8XBN1; UNIPARC:UPI00000D058A; GB:BA000007; PIDN:BA036702.1
A:Experimental source: strain O157:H7, substrain RIMD 0509952
C:Genetics:
A:Gene: ECS3279

Query Match 48.9%; Score 46; DB 2; Length 113;
Best Local Similarity 45.0%; Pred. No. 2.9;
Matches 9; Conservative 5; Mismatches 6; Indels 0; Gaps 0;

QY 1 QAQVNTIQGVSAALQTQAE 20
|||:|||||:|:
DB 65 QQQITLLETQLAQLRQQA 84
|||:|||||:|:

```
RESULT 8
A:Accession: A85883
A:Title: Hypothetical protein Z3672 [imported] - Escherichia coli (strain O157:H7, substrain EDL958)
C:Species: Escherichia coli
C>Date: 16-Feb-2001 #sequence_revision 16-Feb-2001 #text_change 09-Jul-2004
C:Accession: A85883
C:Perna, N.T.; Plunkett III, G.; Burland, V.; Mau, B.; Glasner, J.D.; Rose, D.J.; Mayhew, M.W.;
Miller, L.; Grobeck, E.J.; Davis, N.W.; Lim, A.; Dimalanta, E.; Potamousis, K.; Apodaca, J.;
Nature 409, 529-533, 2001
A:Title: Genome sequence of enterohemorrhagic Escherichia coli O157:H7.
A:Reference number: A85480; MUID:21074935; PMID:11206551
A:Accession: A85883
A>Status: Preliminary
A:Molecule type: DNA
A:Residues: 1-113 <STO>
A:Cross-references: UNIPROT:O8XEN1; UNIPARC:UPI00000D058A; GB:AE005174; NID:g12516779; F
A:Experimental source: strain O157:H7, substrain EDL933
C:Genetics:
A:Gene: Z3672

Query Match 48.9%; Score 46; DB 2; Length 113;
Best Local Similarity 45.0%; Pred. No. 2.9;
Matches 9; Conservative 5; Mismatches 6; Indels 0; Gaps 0;

QY 1 QAVNTIQGVSAALQTQAE 20
DB 65 QQQITLLETQLAQLORQQAE 84

RESULT 9
B83167
A:Title: Hypothetical protein PA3829 [imported] - Pseudomonas aeruginosa (strain PA01)
C:Species: Pseudomonas aeruginosa
C>Date: 15-Sep-2000 #sequence_revision 15-Sep-2000 #text_change 09-Jul-2004
C:Accession: B83167
C:Stover, C.K.; Pham, X.Q.; Erwin, A.L.; Mizoguchi, S.D.; Warren, P.; Hickey, M.J.; B
adman, S.; Yuan, Y.; Brody, L.L.; Coulter, S.N.; Folger, K.R.; Kas, A.; Harbig, K.; Lim,
; Lory, S.; Olson, M.V.
Nature 406, 959-964, 2000
A:Title: Complete genome sequence of Pseudomonas aeruginosa PA01, an opportunistic patho
A:Reference number: A82950; MUID:20437337; PMID:10984043
A:Accession: B83167
A>Status: Preliminary
A:Molecule type: DNA
A:Residues: 1-307 <STO>
A:Cross-references: UNIPROT:O9HXH3; UNIPARC:UPI00000C5B2E; GB:AE004800; GB:AE004091; NID
A:Experimental source: strain PA01
C:Genetics:
A:Gene: PA3829

Query Match 48.9%; Score 46; DB 2; Length 307;
Best Local Similarity 52.6%; Pred. No. 8.3;
Matches 10; Conservative 3; Mismatches 6; Indels 0; Gaps 0;

QY 2 AQVNTIQGVSAALQTQAE 20
DB 132 AQVPTVSGHASALRRTRAE 150

RESULT 10
AH2359
A:Title: Hypothetical protein all4432 [imported] - Nostoc sp. (strain PCC 7120)
C:Species: Nostoc sp. PCC 7120
A:Note: Nostoc sp. strain PCC 7120 is a synonym of Anabaena sp. strain PCC 7120
C>Date: 14-Dec-2001 #sequence_revision 14-Dec-2001 #text_change 09-Jul-2004
C:Accession: AH2359
C:Kazuko, T.; Nakamura, Y.; Wolk, C.P.; Kuritz, T.; Sasamoto, S.; Watanabe, A.; Iriguchi
Nakazaki, N.; Shimpou, M.; Sugimoto, M.; Takazawa, M.; Yamada, M.; Yasuda, M.; Tabata, S
DNA Res. 8, 205-213, 2001
A:Title: Complete Genomic Sequence of the Filamentous Nitrogen-fixing Cyanobacterium Ana
A:Reference number: AB1807; MUID:21595285; PMID:11759840
A:Accession: AH2359
A>Status: Preliminary
```

```
A:Molecule type: DNA
A:Residues: 1-734 <KUR>
A:Cross-references: UNIPROT:O8YXK7; UNIPARC:UPI00000CEB17; GB:BA000019; PIDN:BA076131.1.
A:Experimental source: strain PCC 7120
C:Genetics:
A:Gene: all4432

Query Match 48.9%; Score 46; DB 2; Length 734;
Best Local Similarity 47.1%; Pred. No. 21;
Matches 8; Conservative 6; Mismatches 3; Indels 0; Gaps 0;

QY 1 QAVNTIQGVSAALQTQ 17
DB 230 QAVLNDVGRLEQLRTQ 246

RESULT 11
S71477
A:Title: homeotic protein, ovule-specific - Phalaenopsis sp.
C:Species: Phalaenopsis sp.
C>Date: 19-Mar-1997 #sequence_revision 19-Mar-1997 #text_change 05-Oct-2004
C:Accession: S71477
R:Nadeau, J.A.; Zhang, X.; Li, J.; O'Neill, S.D.
submitted to the EMBL Data Library, August 1995
A:Description: Molecular characterization of ovule development: Isolation of stage- and
A:Reference number: S71477
A:Accession: S71477
A:Molecule type: mRNA
A:Residues: 1-768 <NAD>
A:Cross-references: UNIPROT:O40988; UNIPARC:UPI00000A5978; EMBL:U34743; NID:g1173621; P
A:Experimental source: hybrid SM9108; tissue type ovule
C:Keywords: DNA binding; homeobox; nucleus; transcription regulation
F:90-146/Domain: homeobox homology <HOX>

Query Match 48.9%; Score 46; DB 2; Length 768;
Best Local Similarity 52.6%; Pred. No. 22;
Matches 10; Conservative 3; Mismatches 6; Indels 0; Gaps 0;

QY 2 AQVNTIQGVSAALQTQAE 20
DB 26 AQINILEGQLPLQLQLAE 44

RESULT 12
T46242
A:Title: kinesin-like protein - Arabidopsis thaliana
N:Alternate names: protein T9C5.240
C:Species: Arabidopsis thaliana (mouse-ear cress)
C>Date: 04-Feb-2000 #sequence_revision 04-Feb-2000 #text_change 09-Jul-2004
C:Accession: T46242
R:Rieger, M.; Gabel, C.; Mueller-Auer, S.; Schaefer, M.; Zipp, M.; Mewes, H.W.; Lemcke,
submitted to the Protein Sequence Database, December 1999
A:Reference number: Z23026
A:Accession: T46242
A>Status: Preliminary
A:Molecule type: DNA
A:Residues: 1-813 <RIE>
A:Cross-references: UNIPROT:O9SCV4; UNIPARC:UPI00000A365C; EMBL:AL132964
A:Experimental source: cultivar Columbia; BAC clone T9C5
C:Genetics:
A:Map position: 3
A:Introns: 16/3; 40/3; 79/3; 108/2; 156/3; 188/3; 219/2; 300/2; 350/3; 371/3; 449/3; 48
A:Note: T9C5.240

Query Match 48.9%; Score 46; DB 2; Length 813;
Best Local Similarity 45.0%; Pred. No. 23;
Matches 9; Conservative 5; Mismatches 6; Indels 0; Gaps 0;

QY 1 QAVNTIQGVSAALQTQAE 20
DB 364 QRMINDLQSEVSLQTQLAE 383
```

RESULT 13

H83558
 hypothetical protein PA0697 [imported] - Pseudomonas aeruginosa (strain PA01)
 C:Species: Pseudomonas aeruginosa
 C>Date: 15-Sep-2000 #sequence_revision 15-Sep-2000 #text_change 09-Jul-2004
 C:Accession: H83558
 R:Stover, C.K.; Pham, X.Q.; Erwin, A.L.; Mizoguchi, S.D.; Warriner, P.; Hickey, M.J.; Boman, S.; Yuan, Y.; Brody, L.L.; Coulter, S.N.; Folger, K.R.; Kas, A.; Larbig, K.; Lim, J.; Lory, S.; Olson, M.V.
 Nature 406, 959-964, 2000
 A:Title: Complete genome sequence of Pseudomonas aeruginosa PA01, an opportunistic pathogen
 A:Reference number: AB2950; MUID:20437337; PMID:10984043
 A:Accession: H83558
 A>Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-220 <STO>
 A:Cross-references: UNIPROT:Q9ISM9; UNIPARC:UPI00000C5108; GB:AE004505; GB:AE004091; NID:10984043
 A:Experimental source: strain PA01
 C:Genetics:
 A:Gene: PA0697

Query Match 47.9%; Score 45; DB 2; Length 220;
 Best Local Similarity 45.0%; Pred. No. 8.5;
 Matches 9; Conservative 6; Mismatches 5; Indels 0; Gaps 0;

QY 1 QAQVNTIQGVSAQTQQA 20
 :||: :|: |||||:
 DB 39 RAQLSTTQQLQALQTEQA 58

RESULT 14

AE2304
 hypothetical protein alr3988 [imported] - Nostoc sp. (strain PCC 7120)
 C:Species: Nostoc sp. PCC 7120
 A:Note: Nostoc sp. strain PCC 7120 is a synonym of Anabaena sp. strain PCC 7120
 C>Date: 14-Dec-2001 #sequence_revision 14-Dec-2001 #text_change 09-Jul-2004
 C:Accession: AE2304
 R:Kaneko, T.; Nakamura, Y.; Wolk, C.P.; Kuritz, T.; Sasamoto, S.; Watanabe, A.; Iriguchi, N.; Nakazaki, N.; Shimpo, S.; Sugimoto, M.; Takazawa, M.; Yamada, M.; Tabata, S.
 DNA Res. 8, 205-213, 2001
 A:Title: Complete Genomic Sequence of the Filamentous Nitrogen-fixing Cyanobacterium Anabaena PCC 7120
 A:Reference number: AB1807; MUID:21595285; PMID:11759840
 A:Accession: AE2304
 A>Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-1008 <KUR>
 A:Cross-references: UNIPROT:Q8YQ48; UNIPARC:UPI00000CE995; GB:BA000019; PIDN:BA075687.1
 A:Experimental source: strain PCC 7120
 C:Genetics:
 A:Gene: alr3988

Query Match 47.9%; Score 45; DB 2; Length 1008;
 Best Local Similarity 40.0%; Pred. No. 42;
 Matches 8; Conservative 6; Mismatches 6; Indels 0; Gaps 0;

QY 1 QAQVNTIQGVSAQTQQA 20
 :||: :|: |||||:
 DB 333 ERQLQVQVQLEALQEQE 352

RESULT 15

T13718
 pollux gene protein - fruit fly (Drosophila melanogaster)
 C:Species: Drosophila melanogaster
 C>Date: 13-Aug-1999 #sequence_revision 13-Aug-1999 #text_change 09-Jul-2004
 C:Accession: T13718
 R:Xu, X.Z.; Wes, P.D.; Chen, H.; Li, H.S.; Yu, M.; Morgan, S.; Liu, Y.; Montell, C.
 J. Biol. Chem. 273, 31297, 1998
 A:Title: Retinal targets for calmodulin include proteins implicated in synaptic transmission
 A:Reference number: Z17709; MUID:99030403; PMID:9813038
 A:Accession: T13718
 A>Status: preliminary; translated from GB/EMBL/DD87
 A:Molecule type: mRNA

A:Residues: 1-1379 <XUX>
 A:Cross-references: UNIPROT:Q9TY11; UNIPARC:UPI0000084062; EMBL:Y17919; NID:G3893102; PI
 C:Genetics:
 A:Gene: pollux
 A:Cross-references: FlyBase:FBgn0004879

Query Match 47.9%; Score 45; DB 2; Length 1379;
 Best Local Similarity 53.3%; Pred. No. 59;
 Matches 8; Conservative 4; Mismatches 3; Indels 0; Gaps 0;

QY 1 QAQVNTIQGVSAQLQ 15
 :||: :|: |||||:
 DB 1055 QAQITTLQSQVQSLE 1069

Search completed: January 25, 2006, 19:15:19
 Job time : 9.1 secs

OS Streptococcus mutans.
 OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
 OC Streptococcus.
 OX NCBI_TaxID=1309;
 [1]
 RN NUCLEOTIDE SEQUENCE.
 RC STRAIN=GS-5;
 RA Jin S., Duncan M.J., Taubman M.A., Smith D.J.;
 RT "Cloning of the gbpB gene from Streptococcus mutans.";
 RL J. Dent. Res. 79:224-224 (2000).
 [2]
 RN NUCLEOTIDE SEQUENCE.
 RC STRAIN=GS-5;
 RX MEDLINE=21481971; PubMed=11598068;
 RA Mattos-Graner R.O., Jin S., King W.F., Chen T., Smith D.J.,
 Duncan M.J.;
 RT "Cloning of the Streptococcus mutans gene encoding glucan binding
 protein B and analysis of genetic diversity and protein production in
 clinical isolates.";
 RL Infect. Immun. 69:6931-6941 (2001).
 DR EMBL; AY046410; AAK94500.1; -; Genomic_DNA.
 DR InterPro; IPR007921; CHAP.
 DR InterPro; IPR009148; Siba.
 DR Pfam; PF05257; CHAP; 1.
 DR PRINTS; PR01852; SIBAPROTEIN.
 DR PROSITE; PS50911; CHAP; 1.
 SQ SEQUENCE 431 AA; 44620 MW; 464FE3B563FB7E51 CRC64;

Query Match 100.0%; Score 94; DB 2; Length 431;
 Best Local Similarity 100.0%; Pred. No. 2.1e-06;
 Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 QAQVNTTIQGVSAALQTQAAE 20
 |||||
 DB 52 QAQVNTTIQGVSAALQTQAAE 71

RESULT 3
 Q9AG98 STRMU
 ID Q9AG98 STRMU PRELIMINARY; PRT; 431 AA.
 AC Q9AG98;
 DT 01-JUN-2001 (TrEMBLrel. 17, Created)
 DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
 DT 01-FEB-2005 (TrEMBLrel. 29, Last annotation update)
 DE Immunodominant glycoprotein IDG-60 (Glucan-binding protein B).
 GN Name=saga;
 OS Streptococcus mutans.
 OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
 OC Streptococcus.
 OX NCBI_TaxID=1309;
 [1]
 RN NUCLEOTIDE SEQUENCE.
 RC STRAIN=GS-5;
 RX MEDLINE=21153617; PubMed=11254612;
 RX DOI=10.1128/IAI.69.4.2493-2501.2001;
 RA Chai J.S., Lee Y.Y., Huang P.T., Chen J.Y.;
 RT "Identification of stress-responsive genes in Streptococcus mutans by
 differential display reverse transcription-PCR.";
 RL Infect. Immun. 69:2493-2501 (2001).
 [2]
 RN NUCLEOTIDE SEQUENCE.
 RC STRAIN=GS-5;
 RX MEDLINE=21481977; PubMed=11598074;
 RX DOI=10.1128/IAI.69.11.6987-6998.2001;
 RA Chai J.S., Chang L.Y., Shun C.T., Chang Y.Y., Chen J.Y.;
 RT "A 60-kilodalton immunodominant glycoprotein is essential for cell
 wall integrity and the maintenance of cell shape in Streptococcus
 mutans.";
 RL Infect. Immun. 69:6987-6998 (2001).
 [3]
 RN NUCLEOTIDE SEQUENCE.
 RC STRAIN=GS-5;

RA Chia J.-S., Chang L.-Y., Lee Y.-Y., Chen J.-Y.;
 RL Submitted (NOV-2001) to the EMBL/GenBank/DBJ databases.
 [4]
 RN NUCLEOTIDE SEQUENCE.
 RC STRAIN=3VP4;
 RA Jin S., Duncan M.J., Taubman M.A., Smith D.J.;
 RT "Cloning of the gbpB gene from Streptococcus mutans.";
 RL J. Dent. Res. 79:224-224 (2000).
 [5]
 RN NUCLEOTIDE SEQUENCE.
 RC STRAIN=3VP4;
 RX MEDLINE=21481971; PubMed=11598068;
 RX DOI=10.1128/IAI.69.11.6931-6941.2001;
 RA Mattos-Graner R.O., Jin S., King W.F., Chen T., Smith D.J.,
 Duncan M.J.;
 RT "Cloning of the Streptococcus mutans gene encoding glucan binding
 protein B and analysis of genetic diversity and protein production in
 clinical isolates.";
 RL Infect. Immun. 69:6931-6941 (2001).
 DR EMBL; AF338445; AAK08104.1; -; Genomic_DNA.
 DR EMBL; AY046411; AAK94501.1; -; Genomic_DNA.
 DR InterPro; IPR007921; CHAP.
 DR InterPro; IPR009148; Siba.
 DR Pfam; PF05257; CHAP; 1.
 DR PRINTS; PR01852; SIBAPROTEIN.
 DR PROSITE; PS50911; CHAP; 1.
 SQ SEQUENCE 431 AA; 44592 MW; 3EBE21FC5E47232E CRC64;

Query Match 100.0%; Score 94; DB 2; Length 431;
 Best Local Similarity 100.0%; Pred. No. 2.1e-06;
 Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 QAQVNTTIQGVSAALQTQAAE 20
 |||||
 DB 52 QAQVNTTIQGVSAALQTQAAE 71

RESULT 4
 Q8DWM3 STRMU
 ID Q8DWM3 STRMU PRELIMINARY; PRT; 431 AA.
 AC Q8DWM3;
 DT 01-MAR-2003 (TrEMBLrel. 23, Created)
 DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
 DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
 DE Putative secreted antigen GbpB/Saga; putative peptidoglycan
 DE hydrolase.
 GN Name=gbpB; OrderedLocusNames=SMU_22;
 OS Streptococcus mutans.
 OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
 OC Streptococcus.
 OX NCBI_TaxID=1309;
 [1]
 RN NUCLEOTIDE SEQUENCE.
 RC STRAIN=UA159 / ATCC 700610 / Serotype c;
 RX MEDLINE=22295063; PubMed=12397186; DOI=10.1073/pnas.172501299;
 RA Ajdic D.J., McShan W.M., McLaughlin R.E., Savic G., Chang J.,
 RA Carson M.B., Primeaux C., Tian R., Kenton S., Jia H.G., Lin S.P.,
 RA Qian Y., Li S., Zhu H., Najjar F.Z., Lai H., White J., Roe B.A.,
 RA Ferretti J.J.;
 RT "Genome sequence of Streptococcus mutans UA159, a cariogenic dental
 pathogen.";
 RL Proc. Natl. Acad. Sci. U.S.A. 99:14434-14439 (2002).
 DR EMBL; AE014855; AAN57811.1; -; Genomic_DNA.
 DR InterPro; IPR007921; CHAP.
 DR InterPro; IPR009148; Siba.
 DR Pfam; PF05257; CHAP; 1.
 DR PRINTS; PR01852; SIBAPROTEIN.
 DR PROSITE; PS50911; CHAP; 1.
 KW Complete proteome.
 SQ SEQUENCE 431 AA; 44620 MW; 2D1CA685248CCD3E CRC64;

Query Match 100.0%; Score 94; DB 2; Length 431;
 Best Local Similarity 100.0%; Pred. No. 2.1e-06;

Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 QAQVNTIQGVSAALQTQAAE 20
 |||||
 Db 52 QAQVNTIQGVSAALQTQAAE 71

RESULT 5

Q938V1_STRMU
 ID Q938V1_STRMU PRELIMINARY; PRT; 432 AA.
 AC Q938V1;
 DT 01-DEC-2001 (TReMBLrel. 19, Created)
 DT 01-DEC-2001 (TReMBLrel. 19, Last sequence update)
 DT 01-WAR-2004 (TReMBLrel. 26, Last annotation update)
 DE Glucan-binding protein B.
 OS Streptococcus mutans.
 OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
 OC Streptococcus.
 OX NCBI_TaxID=1309;
 RN [1]
 RP NUCLEOTIDE SEQUENCE.
 RC STRAIN=3SN1;
 RA Jin S., Duncan M.J., Taubman M.A., Smith D.J.;
 RT "Cloning of the gbpB gene from Streptococcus mutans.";
 RL J. Dent. Res. 79:224-224 (2000).
 RN [2]

NUCLEOTIDE SEQUENCE.

RC STRAIN=3SN1;
 RX MEDLINE=21481971; PubMed=11598068;
 RX DOI=10.1128/IAI.69.11.6931-6941.2001;
 RA Mattos-Graner R.O., Jin S., King W.F., Chen T., Smith D.J.,
 RA Duncan M.J.;
 RT "Cloning of the Streptococcus mutans gene encoding glucan binding
 RT protein B and analysis of genetic diversity and protein production in
 RT clinical isolates.";
 RL Infect. Immun. 69:6931-6941 (2001).
 DR EMBL; AY046413; AAK94503.1; -; Genomic_DNA.
 DR InterPro; IPR007921; CHAP.
 DR Pfam; PF05257; CHAP; 1.
 DR PRINTS; PRO1852; SIBAPROTEIN.
 DR PROSITE; PS50911; CHAP; 1.
 SQ SEQUENCE 432 AA; 44652 MW; 3F88ECB9A1F3BE4F CRC64;

Query Match 100.0%; Score 94; DB 2; Length 432;

Best Local Similarity 100.0%; Pred. No. 2.1e-06;
 Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 QAQVNTIQGVSAALQTQAAE 20
 |||||
 Db 52 QAQVNTIQGVSAALQTQAAE 71

RESULT 6

Q938V2_STRMU
 ID Q938V2_STRMU PRELIMINARY; PRT; 432 AA.
 AC Q938V2;
 DT 01-DEC-2001 (TReMBLrel. 19, Created)
 DT 01-DEC-2001 (TReMBLrel. 19, Last sequence update)
 DT 01-WAR-2004 (TReMBLrel. 26, Last annotation update)
 DE Glucan-binding protein B.
 OS Streptococcus mutans.
 OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
 OC Streptococcus.
 OX NCBI_TaxID=1309;
 RN [1]
 RP NUCLEOTIDE SEQUENCE.
 RC STRAIN=15JP2;
 RA Jin S., Duncan M.J., Taubman M.A., Smith D.J.;
 RT "Cloning of the gbpB gene from Streptococcus mutans.";
 RL J. Dent. Res. 79:224-224 (2000).
 RN [2]

NUCLEOTIDE SEQUENCE.

RC STRAIN=15JP2;
 RX MEDLINE=21481971; PubMed=11598068;
 RX DOI=10.1128/IAI.69.11.6931-6941.2001;
 RA Mattos-Graner R.O., Jin S., King W.F., Chen T., Smith D.J.,
 RA Duncan M.J.;
 RT "Cloning of the Streptococcus mutans gene encoding glucan binding
 RT protein B and analysis of genetic diversity and protein production in
 RT clinical isolates.";
 RL Infect. Immun. 69:6931-6941 (2001).
 DR EMBL; AY046413; AAK94503.1; -; Genomic_DNA.
 DR InterPro; IPR007921; CHAP.
 DR Pfam; PF05257; CHAP; 1.
 DR PRINTS; PRO1852; SIBAPROTEIN.
 DR PROSITE; PS50911; CHAP; 1.
 SQ SEQUENCE 432 AA; 44652 MW; 3F88ECB9A1F3BE4F CRC64;

RC STRAIN=15JP2;
 RX MEDLINE=21481971; PubMed=11598068;
 RX DOI=10.1128/IAI.69.11.6931-6941.2001;
 RA Mattos-Graner R.O., Jin S., King W.F., Chen T., Smith D.J.,
 RA Duncan M.J.;
 RT "Cloning of the Streptococcus mutans gene encoding glucan binding
 RT protein B and analysis of genetic diversity and protein production in
 RT clinical isolates.";
 RL Infect. Immun. 69:6931-6941 (2001).
 DR EMBL; AY046413; AAK94503.1; -; Genomic_DNA.
 DR InterPro; IPR007921; CHAP.
 DR Pfam; PF05257; CHAP; 1.
 DR PRINTS; PRO1852; SIBAPROTEIN.
 DR PROSITE; PS50911; CHAP; 1.
 SQ SEQUENCE 432 AA; 44648 MW; E769B2504AE50E9 CRC64;

Query Match 100.0%; Score 94; DB 2; Length 432;

Best Local Similarity 100.0%; Pred. No. 2.1e-06;
 Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 QAQVNTIQGVSAALQTQAAE 20
 |||||
 Db 52 QAQVNTIQGVSAALQTQAAE 71

RESULT 7

Q9ZAS7_STRMU
 ID Q9ZAS7_STRMU PRELIMINARY; PRT; 211 AA.
 AC Q9ZAS7;
 DT 01-MAY-1999 (TReMBLrel. 10, Created)
 DT 01-MAY-1999 (TReMBLrel. 10, Last sequence update)
 DT 01-OCT-2002 (TReMBLrel. 22, Last annotation update)
 DE Putative secreted protein (Fragment).
 OS Streptococcus mutans.
 OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
 OC Streptococcus.
 OX NCBI_TaxID=1309;
 RN [1]
 RP NUCLEOTIDE SEQUENCE.
 RC STRAIN=GS-5 Kuramitsu;
 RA Peruzzi F., Piggot P.J., Daneo-Moore L.;
 RL Submitted (NOV-1996) to the EMBL/GenBank/DBJ databases.
 DR EMBL; U78607; AAD00288.1; -; Genomic_DNA.
 FT NON TER 211 211
 SQ SEQUENCE 211 AA; 22803 MW; 4ACE331159CFAPFC6 CRC64;

Query Match 74.5%; Score 70; DB 2; Length 211;

Best Local Similarity 100.0%; Pred. No. 0.0069;
 Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 6 TIQGVSAALQTQAAE 20
 |||||
 Db 57 TIQGVSAALQTQAAE 71

RESULT 8

Q5M6K4_STRT2
 ID Q5M6K4_STRT2 PRELIMINARY; PRT; 474 AA.
 AC Q5M6K4;
 DT 01-FEB-2005 (TReMBLrel. 29, Created)
 DT 01-FEB-2005 (TReMBLrel. 29, Last sequence update)
 DT 13-SEP-2005 (TReMBLrel. 31, Last annotation update)
 DE Glucan binding protein (PcsB).
 GN Name=pcsB; OrderedLocNames=stu0022;
 OS Streptococcus thermophilus (strain ATCC BAA-250 / LMG 18311).
 OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
 OC Streptococcus.
 OX NCBI_TaxID=264199;
 RN [1]
 RP NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
 RX PubMed=15543133; DOI=10.1038/nbt1034;
 RA Bolotin A., Quinquis B., Renault P., Sorokin A., Ehrlich S.D.,

```

Db          49 QAOVATIQSVTLRTQKTE 68

RESULT 10
Q9AKA4_STRAG PRELIMINARY; PRT; 447 AA.
ID Q9AKA4;
AC AC
DT 01-JUN-2001 (TrEMBLrel. 17, Created)
DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE PcsB protein precursor.
GN Name=PcsB;
OS Streptococcus agalactiae.
OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
OC Streptococcus.
OX NCBI_TaxID=1311;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=6313;
RX MEDLINE=21101799; PubMed=11157929;
RA DOI=10.1128/JB.183.4.1175-1183.2001;
RA Reinscheid D.J., Gottechalk B., Schubert A., Eikmanns B.J.,
RA Chhatwal G.S.;
RT "Identification and molecular analysis of PcsB, a protein required for
RT cell wall separation of group B streptococcus.";
RL J. Bacteriol. 183:1175-1183(2001).
DR EMBL; AJ277292; CAC28144.1; - ; Genomic_DNA.
DR InterPro; IPR007921; CHAP.
DR InterPro; IPR009148; Siba.
DR Pfam; PF05257; CHAP; 1.
DR PRINTS; PR01852; SIBAPROTEIN.
DR PROSITE; PS09111; CHAP; 1.
FT SIGNAL.
FT FTCHAIN 1 25 Potential.
FT FTSEQUENCE 26 447 PcsB protein.
SQ SEQUENCE 447 AA; 46681 MW; F4DBI4BOASF62C8 CRC64;

Query Match 62.8%; Score 59; DB 2; Length 447;
Best Local Similarity 60.0%; Pred. No. 0.84;
Matches 12; Conservative 4; Mismatches 4; Indels 0; Gaps 0;

OY 1 QAOVNTIQGVSALETQQAE 20
| | | : | | | : | |
Db 50 QNQVTAIKGQGVALESQQSE 69

RESULT 11
Q8E2H1_STRAS PRELIMINARY; PRT; 447 AA.
ID Q8E2H1;
AC AC
DT 01-MAR-2003 (TrEMBLrel. 23, Created)
DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE PcsB protein.
GN Name=PcsB; OrderedLocusNames=SAG0017;
OS Streptococcus agalactiae (serotype V).
OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
OC Streptococcus.
OX NCBI_TaxID=216466;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=2603 V/R / Serotype V;
RX MEDLINE=22222998; PubMed=12200547; DOI=10.1073/pnas.182380799;
RA Tettelin H., Masignani V., Cieslewicz M.J., Eisen J.A., Peterson S.N.,
RA Wessels L.C., Paulsen I.T., Nelson K.E., Margarit I., Read T.D.,
RA Madoff L.R., Wolf A.M., Beanan M.J., Brinkac L.M., Daugherty S.C.,
RA DeBoy R.T., Durkin A.S., Kolonay J.F., Madupu R., Lewis M.R.,
RA Radune D., Fedorova N.B., Scanlan D., Khouri H.M., Mulligan S.,
RA Carty H.A., Cline R.T., Van Aken S.E., Gill J., Scarselli M., Mora M.,
RA Iacobini E.T., Brettoni C., Galli G., Mariani M., Vegni F., Malone D.,
RA Rinaudo D., Rappuoli R., Telford J.B., Kasper D.L., Grandi G.,
RA Fraser C.M.;
RT "Complete genome sequence and comparative genomic analysis of an

```

RT emerging human pathogen, serotype V Streptococcus agalactiae.";
 RL Proc. Natl. Acad. Sci. U.S.A. 99:12391-12396 (2002).
 DR EMBL; AE014192; AAM98925.1; -; Genomic_DNA.
 DR TIGR; SAG0017; -.

DR InterPro; IPR007921; CHAP.
 DR InterPro; IPR009148; Siba.
 DR Pfam; PF05257; CHAP; 1.
 DR PRINTS; PR01852; SIBAPROTEIN.
 DR PROSITE; PS0911; CHAP; 1.
 KW Complete proteome.
 SQ SEQUENCE 447 AA; 46681 MW; F4DB14B0A5F962C8 CRC64;

Query Match 62.8%; Score 59; DB 2; Length 447;
 Best Local Similarity 60.0%; Pred. No. 0.84;
 Matches 12; Conservative 4; Mismatches 4; Indels 0; Gaps 0;

QY 1 QAQVNTIQGVSAALOTQQA 20
 |||:|||||:|||||:
 Db 50 QNQVTAIKGVGALSSQQSE 69

RESULT 12

Q8E7X9 STRA3
 ID Q8E7X9 STRA3 PRELIMINARY; PRT; 447 AA.
 AC Q8E7X9

DT 01-MAR-2003 (TrEMBLrel. 23, Created)
 DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
 DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
 DE PCSB protein.

GN Name=pcsb; OrderedLocustNames=gbs0016;
 OS Streptococcus agalactiae (serotype III).

OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
 OC Streptococcus.

OX NCBI_TaxID=216495;
 [1]

RN NUCLEOTIDE SEQUENCE.
 RP STRAIN-NEW316 / Serotype III;

RX MEDLINE=2242508; PubMed=12354221;
 RA Glaser P., Rueniok C., Buchrieser C., Chevalier F., Frangeul L.,
 RA Msadek T., Zouine M., Couve E., Lalioui L., Poyart C., Trieu-Cuot P.,
 RA Kunst F.;

RT "Genome sequence of Streptococcus agalactiae, a pathogen causing
 RT invasive neonatal disease.";
 RL Mol. Microbiol. 45:1499-1513 (2002).
 DR EMBL; AL766843; CAD45661.1; -; Genomic_DNA.
 DR SagsList; gbs0016; -.

DR InterPro; IPR007921; CHAP.
 DR InterPro; IPR009148; Siba.
 DR Pfam; PF05257; CHAP; 1.

DR PRINTS; PR01852; SIBAPROTEIN.
 DR PROSITE; PS0911; CHAP; 1.

KW Complete proteome.
 SQ SEQUENCE 447 AA; 46681 MW; F4DB14B0A5F962C8 CRC64;

Query Match 62.8%; Score 59; DB 2; Length 447;
 Best Local Similarity 60.0%; Pred. No. 0.84;
 Matches 12; Conservative 4; Mismatches 4; Indels 0; Gaps 0;

QY 1 QAQVNTIQGVSAALOTQQA 20
 |||:|||||:|||||:
 Db 50 QNQVTAIKGVGALSSQQSE 69

RESULT 13

Q5HUT7 CAMJR

ID Q5HUT7 CAMJR PRELIMINARY; PRT; 142 AA.
 AC Q5HUT7

DT 10-MAY-2005 (TrEMBLrel. 30, Created)
 DT 10-MAY-2005 (TrEMBLrel. 30, Last sequence update)
 DT 10-MAY-2005 (TrEMBLrel. 30, Last annotation update)

DE Hypothetical protein.
 GN OrderedLocustNames=CJ08946;

OS Campylobacter jejuni (strain RM1221).

OC Bacteria; Proteobacteria; Epsilonproteobacteria; Campylobacteriales;
 OC Campylobacteraceae; Campylobacter.

OX NCBI_TaxID=195099;
 [1]

RN NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
 RX PubMed=15660156; DOI=10.1371/journal.pbio.0030015;

RA Fouts D.E., Mongodin E.P., Mandrell R.E., Miller W.G., Raeko D.A.,
 RA Ravel J., Brinkac L.M., DeBoy R.T., Parker C.T., Daugherty S.C.,
 RA Dodson R.J., Durkin A.S., Madupu R., Sullivan S.A., Shetty J.U.,
 RA Ayodeji M.A., Shvartsbeyn A., Schatz M.C., Badger J.H., Fraser C.M.,
 RA Nelson K.E.;

RT "Major structural differences and novel potential virulence mechanisms
 RT from the genomes of multiple Campylobacter species.";
 RL PLOS Biol. 3:72-85 (2005).
 DR EMBL; CP000025; AAW35283.1; -; Genomic_DNA.
 DR TIGR; CJ0946; -.

KW Complete proteome; Hypothetical protein.
 SQ SEQUENCE 142 AA; 15570 MW; 1A2CE595CC5FD19D CRC64;

Query Match 58.5%; Score 55; DB 2; Length 142;
 Best Local Similarity 61.1%; Pred. No. 1.2;
 Matches 11; Conservative 3; Mismatches 4; Indels 0; Gaps 0;

QY 2 AQVNTIQGVSAALOTQQA 19
 |||:|||||:|||||:
 Db 125 AQASTIAQIQELQSQA 142

RESULT 14

Q9PP64 CAMJE

ID Q9PP64 CAMJE PRELIMINARY; PRT; 142 AA.
 AC Q9PP64

DT 01-OCT-2000 (TrEMBLrel. 15, Created)
 DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)

DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
 DE Hypothetical protein Cj0859c.

GN OrderedLocustNames=Cj0859.3;
 OS Campylobacter jejuni.

OC Bacteria; Proteobacteria; Epsilonproteobacteria; Campylobacteriales;
 OC Campylobacteraceae; Campylobacter.

OX NCBI_TaxID=197;
 [1]

RN NUCLEOTIDE SEQUENCE.

RC STRAIN-NCTC 11168;

RX MEDLINE=20150912; PubMed=10688204; DOI=10.1038/35001088;
 RA Parkhill J., Wren B.W., Mungall K.L., Ketley J.M., Churcher C.M.,
 RA Baoham D., Chillingworth T., Davies R.M., Feltwell T., Holroyd S.,
 RA Jagels K., Karlyshev A.V., Moule S., Pallen M.J., Penn C.W.,
 RA Quail M.A., Rajandream M.A., Rutherford K.M., van Vliet A.H.M.,
 RA Whitehead S., Barrell B.G.;

RT "The genome sequence of the food-borne pathogen Campylobacter jejuni
 RT reveals hypervariable sequences.";
 RL Nature 403:665-668 (2000).

DR EMBL; AL139076; CAB73124.1; -; Genomic_DNA.
 DR PIR; C81359; C81359.

KW Complete proteome; Hypothetical protein.
 SQ SEQUENCE 142 AA; 15545 MW; 982CF70B2F2FC073 CRC64;

Query Match 58.5%; Score 55; DB 2; Length 142;
 Best Local Similarity 61.1%; Pred. No. 1.2;
 Matches 11; Conservative 3; Mismatches 4; Indels 0; Gaps 0;

QY 2 AQVNTIQGVSAALOTQQA 19
 |||:|||||:|||||:
 Db 125 AQASTIAQIQELQSQA 142

RESULT 15

Q5LSX9 CHLAB

ID Q5LSX9 CHLAB PRELIMINARY; PRT; 366 AA.
 AC Q5LSX9

DT 01-FEB-2005 (TrEMBLrel. 29, Created)
 DT 01-FEB-2005 (TrEMBLrel. 29, Last sequence update)

DT 01-FEB-2005 (TrEMBLrel. 29, Last annotation update)
 DE Putative inner membrane protein.
 GN OrderedLocusNames=CAB500;
 OS Chlamydomophila abortus.
 OC Bacteria; Chlamydiae; Chlamydiales; Chlamydiaceae; Chlamydomophila.
 OX NCBI_TaxID=83555;
 RN [1]
 RP NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
 RC STRAIN=S26/3;
 RX PubMed=15837807; DOI=10.1101/gr.3684805;
 RA Thomson N.R., Yeats C., Bell K., Hoiden M.T.G., Bentley S.D.,
 RA Livingstone M., Cerdano-Tarraga A.-M., Harris B., Doggett J.,
 RA Ormond D., Mungall K., Clarke K., Feltwell T., Hance Z., Sanders M.,
 RA Quail M.A., Price C., Barrall B.G., Parkhill J., Longbottom D.;
 RT "The Chlamydomophila abortus genome sequence reveals an array of
 RT variable proteins that contribute to interspecies variation.";
 RL Genome Res. 15:629-640(2005).
 DR EMBL; CR848038; CAH63952.1; -; Genomic_DNA.
 KW Complete proteome.
 SQ SEQUENCE 366 AA; 39681 MW; 17C3FA52934F22E0 CRC64;

 Query Match 56.4%; Score 53; DB 2; Length 366;
 Best Local Similarity 61.1%; Pred. NO. 6.3;
 Matches 11; Conservative 3; Mismatches 4; Indels 0; Gaps 0;

 Qy 3 QVNTIQGVSA LQTQAE 20
 :||:|||||
 Db 120 EVNRLQGEVSASQAQLAE 137

Search completed: January 25, 2006, 19:13:41
 Job time : 71.3 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2006 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: January 25, 2006, 18:50:05 ; Search time 16.6 Seconds
(without alignments)
99.609 Million cell updates/sec

Title: US-10-797-821-9
Perfect score: 94
Sequence: 1 QAVNTIQGVSAALQTQQA 20

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 572060 seqs, 82675679 residues

Total number of hits satisfying chosen parameters: 572060

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued Patents AA:*
1: /cgn2_6/ptodata/1/iaa/5-COMB.pep:*
2: /cgn2_6/ptodata/1/iaa/6-COMB.pep:*
3: /cgn2_6/ptodata/1/iaa/H-COMB.pep:*
4: /cgn2_6/ptodata/1/iaa/PCITUS-COMB.pep:*
5: /cgn2_6/ptodata/1/iaa/RE-COMB.pep:*
6: /cgn2_6/ptodata/1/iaa/backfiles1.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	52	55.3	210	2 US-09-222-938A-67	Sequence 67, Appl
2	52	55.3	392	2 US-09-583-110-4374	Sequence 4374, Ap
3	52	55.3	399	2 US-09-107-433-3230	Sequence 3230, Ap
4	47	50.0	461	1 US-08-186-222-2	Sequence 2, Appli
5	47	50.0	1024	2 US-09-562-737-45	Sequence 45, Appl
6	46	48.9	722	1 US-09-252-991A-22811	Sequence 22811, A
7	46	48.9	768	1 US-08-560-398-2	Sequence 2, Appli
8	45	47.9	383	2 US-09-252-991A-29621	Sequence 29621, A
9	45	47.9	1024	2 US-09-562-737-43	Sequence 43, Appl
10	45	47.9	3165	1 US-08-459-146-3	Sequence 3, Appli
11	45	47.9	3165	1 US-08-459-065-3	Sequence 3, Appli
12	44	46.8	76	1 US-08-338-558-1	Sequence 1, Appli
13	44	46.8	134	1 US-08-850-910A-45	Sequence 45, Appl
14	44	46.8	134	1 US-08-850-910A-48	Sequence 48, Appl
15	44	46.8	134	2 US-09-508-435A-2	Sequence 2, Appli
16	44	46.8	134	2 US-09-502-517-45	Sequence 45, Appl
17	44	46.8	134	2 US-09-902-517-48	Sequence 48, Appl
18	44	46.8	329	2 US-09-543-681A-5999	Sequence 5999, Ap
19	44	46.8	373	2 US-09-107-532A-7035	Sequence 7035, Ap
20	44	46.8	1024	2 US-09-562-737-41	Sequence 41, Appl
21	44	46.8	1024	2 US-09-562-737-47	Sequence 47, Appl
22	43.5	46.3	386	2 US-09-489-039A-7410	Sequence 7410, Ap
23	43	45.7	218	2 US-09-603-208A-30	Sequence 30, Appl
24	43	45.7	272	2 US-09-252-991A-26554	Sequence 26554, A
25	43	45.7	287	2 US-08-937-271-2	Sequence 2, Appli
26	43	45.7	525	2 US-09-107-532A-5095	Sequence 5095, Ap
27	42.5	45.2	41	2 US-08-737-629-6	Sequence 6, Appli

28	42	44.7	243	2 US-09-252-991A-26233	Sequence 26233, A
29	42	44.7	365	2 US-09-328-352-7412	Sequence 7412, Ap
30	42	44.7	804	2 US-09-252-991A-24051	Sequence 24051, A
31	41.5	44.1	190	2 US-09-538-092-86	Sequence 86, Appl
32	41.5	44.1	487	2 US-09-252-991A-29118	Sequence 29118, A
33	41	43.6	35	2 US-09-270-767-58615	Sequence 58615, A
34	41	43.6	100	2 US-09-370-838-94	Sequence 94, Appl
35	41	43.6	100	2 US-09-854-133-94	Sequence 94, Appl
36	41	43.6	234	2 US-09-270-767-43273	Sequence 43273, A
37	41	43.6	263	2 US-09-252-991A-24295	Sequence 24295, A
38	41	43.6	292	2 US-09-134-001C-3479	Sequence 3479, Ap
39	41	43.6	549	2 US-09-134-001C-2973	Sequence 2973, Ap
40	41	43.6	593	2 US-09-538-092-919	Sequence 919, App
41	41	43.6	2101	1 US-08-466-390-4	Sequence 4, Appli
42	41	43.6	2101	1 US-08-470-950-4	Sequence 4, Appli
43	41	43.6	2101	1 US-08-467-781-4	Sequence 4, Appli
44	41	43.6	2101	1 US-08-195-487-4	Sequence 4, Appli
45	41	43.6	2101	1 US-08-483-924-4	Sequence 4, Appli

ALIGNMENTS

RESULT 1
US-09-222-938A-67
; Sequence 67, Application US/09222938A
; Patent No. 6437108
; GENERAL INFORMATION:
; APPLICANT: Youngman, Philip
; APPLICANT: Fritz, Christian
; APPLICANT: Murphy, Christopher
; APPLICANT: Guzman, Luz-Maria
; TITLE OF INVENTION: ESSENTIAL BACTERIAL GENES AND THEIR USE
; FILE REFERENCE: 07334/060001
; CURRENT APPLICATION NUMBER: US/09/222,938A
; CURRENT FILING DATE: 1998-12-30
; NUMBER OF SEQ ID NOS: 102
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 67
; TYPE: PRT
; ORGANISM: Streptococcus pneumoniae
US-09-222-938A-67

Query Match 55.3%; Score 52; DB 2; Length 210;
Best Local Similarity 57.9%; Pred. No. 0.48;
Matches 11; Conservative 4; Mismatches 4; Indels 0; Gaps 0;

QY 1 QAVNTIQGVSAALQTQQA 19
|||:|||||:|:
Db 52 QKQVDIQGVSAIQAEQS 70

RESULT 2
US-09-583-110-4374
; Sequence 4374, Application US/09583110
; Patent No. 6699703
; GENERAL INFORMATION:
; APPLICANT: Lynn Doucette-Stamm et al.
; TITLE OF INVENTION: Nucleic Acid and Amino Acid Sequences Relating to Streptococcus
; FILE REFERENCE: PATH00-07A
; CURRENT APPLICATION NUMBER: US/09/583,110
; CURRENT FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: US 09/107,433
; PRIOR FILING DATE: 1998-06-30
; PRIOR APPLICATION NUMBER: US 60/085,131
; PRIOR FILING DATE: 1998-05-12
; PRIOR APPLICATION NUMBER: US 60/051,553
; PRIOR FILING DATE: 1997-07-02
; NUMBER OF SEQ ID NOS: 5322
; SEQ ID NO 4374
; LENGTH: 392

; TYPE: PRT
; ORGANISM: Streptococcus pneumoniae
US-09-583-110-4374

Query Match 55.3%; Score 52; DB 2; Length 392;
Best Local Similarity 57.9%; Pred. No. 1;
Matches 11; Conservative 4; Mismatches 4; Indels 0; Gaps 0;

QY 1 QAOVNTIQGVSAALQTQQA 19
Db 52 QKQVDQIQGVSAIQAEQS 70

RESULT 3

US-09-107-433-3230
; Sequence 3230, Application US/09107433
; Patent No. 6800744

; GENERAL INFORMATION:
; APPLICANT: Lynn A Doucette-Stamm and David Bush
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID
; SEQUENCES RELATING TO STREPTOCOCCUS PNEUMONIAE
; THERAPEUTICS

; NUMBER OF SEQUENCES: 5206
; CORRESPONDENCE ADDRESS:
; ADDRESS: GENOME THERAPEUTICS CORPORATION
; STREET: 100 Beaver Street
; CITY: Waltham
; STATE: Massachusetts
; COUNTRY: USA
; ZIP: 02354

; COMPUTER READABLE FORM:
; MEDIUM TYPE: CD-ROM ISO9660
; COMPUTER: <Unknown>
; OPERATING SYSTEM: <Unknown>
; SOFTWARE: <Unknown>

; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/107,433
; FILING DATE: 30-Jun-1998

; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 60/ 085131
; FILING DATE: May 12, 1998
; APPLICATION NUMBER: 60/051553
; FILING DATE: July 2, 1997

; ATTORNEY/AGENT INFORMATION:
; NAME: Ariniello, Pamela Deneka
; REGISTRATION NUMBER: 40,489
; REFERENCE/DOCKET NUMBER: GTC-011
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (781)893-5007
; TELEFAX: (781)893-8277

; INFORMATION FOR SEQ ID NO: 3230:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 399 amino acids
; TYPE: amino acid
; TOPOLOGY: linear

; MOLECULE TYPE: protein
; HYPOTHETICAL: YES
; ORIGINAL SOURCE:
; ORGANISM: Streptococcus pneumoniae
; FEATURE:

; NAME/KEY: misc_feature
; LOCATION: (B) LOCATION 1...399
; SEQUENCE DESCRIPTION: SEQ ID NO: 3230:
US-09-107-433-3230

Query Match 55.3%; Score 52; DB 2; Length 399;
Best Local Similarity 57.9%; Pred. No. 1;
Matches 11; Conservative 4; Mismatches 4; Indels 0; Gaps 0;

QY 1 QAOVNTIQGVSAALQTQQA 19
Db 59 QKQVDQIQGVSAIQAEQS 77

RESULT 4

US-08-186-222-2
; Sequence 2, Application US/08186222
; Patent No. 5559007

; GENERAL INFORMATION:
; APPLICANT: Suri, Bruno
; TITLE OF INVENTION: Bacterial Vectors
; NUMBER OF SEQUENCES: 4

; CORRESPONDENCE ADDRESS:
; ADDRESSEE: CIBA-GEIGY Corporation
; STREET: 7 Skyline Drive
; CITY: Hawthorne
; STATE: New York

; COUNTRY: USA
; ZIP: 10532
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk

; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/186,222
; FILING DATE:

; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/672,205
; FILING DATE: 19-MAR-1991

; APPLICATION NUMBER: GB 9006400.7
; FILING DATE: 22-MAR-1990
; ATTORNEY/AGENT INFORMATION:
; NAME: Villamizar, JoAnn

; REGISTRATION NUMBER: 30,598
; REFERENCE/DOCKET NUMBER: 4-17994/A
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (914)785-7121
; TELEFAX: (914)347-5769

; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 461 amino acids
; TYPE: amino acid
; TOPOLOGY: linear

; MOLECULE TYPE: protein
US-08-186-222-2

Query Match 50.0%; Score 47; DB 1; Length 461;
Best Local Similarity 50.0%; Pred. No. 8.5;
Matches 9; Conservative 6; Mismatches 3; Indels 0; Gaps 0;

QY 1 QAOVNTIQGVSAALQTQQ 18
Db 51 QAOVDSLOSQVDSLQQKQ 68

RESULT 5

US-09-562-737-45

; Sequence 45, Application US/09562737
; Patent No. 6428967

; GENERAL INFORMATION:
; APPLICANT: Herz, Joachim
; APPLICANT: Gotthardt, Michael
; TITLE OF INVENTION: Ldl Receptor Signaling Pathways
; FILE REFERENCE: UTSW0708

; CURRENT APPLICATION NUMBER: US/09/562,737
; CURRENT FILING DATE: 2000-05-01
; NUMBER OF SEQ ID NOS: 132

; SOFTWARE: PatentIn Ver. 2.1

; SEQ ID NO 45

; LENGTH: 1024

; TYPE: PRT

; ORGANISM: Artificial Sequence

; FEATURE:

REGISTRATION NUMBER: 34,774
REFERENCE/DOCKET NUMBER: 0230

US-09-562-737-43

Patent
GEN
A

; APPLICATION NUMBER: US/08/338,558
; FILING DATE: 21-NOV-1994
; CLASSIFICATION: 436
; ATTORNEY/AGENT INFORMATION:
; NAME: Fanucci, Allan A.
; REGISTRATION NUMBER: 30,256
; REFERENCE/DOCKET NUMBER: 8080-010
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 790-9090
; TELEFAX: (212) 869-9741/8864
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 76 amino acids
; TYPE: amino acid
; TOPOLOGY: unknown
; MOLECULE TYPE: protein
US-08-338-558-1

Query Match 46.8%; Score 44; DB 1; Length 76;
Best Local Similarity 50.0%; Pred. No. 3.2;
Matches 9; Conservative 4; Mismatches 5; Indels 0; Gaps 0;

QY 1 QAOVNTIQGVSAALQTOQ 18
DB 18 QEQRNHLOKLSLQVEQ 35

RESULT 13
US-08-850-910A-45
; Sequence 45, Application US/08850910A
; Patent No. 5948761
; GENERAL INFORMATION:
; APPLICANT: SEILHAMER, J. J.
; APPLICANT: LEWICKI, J.
; APPLICANT: SCARBOROUGH, R. M.
; TITLE OF INVENTION: RECOMBINANT TECHNIQUES FOR
; TITLE OF INVENTION: PRODUCTION OF BRAIN NATRIUETIC PEPTIDE
; NUMBER OF SEQUENCES: 50
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: MORRISON & FOERSTER, LLP
; STREET: 2000 Pennsylvania Avenue, NW, Suite 5500
; CITY: Washington
; STATE: DC
; COUNTRY: USA
; ZIP: 20006-1888
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: Windows
; SOFTWARE: FastSeq for Windows Version 2.0b
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/850,910A
; FILING DATE: 05-MAY-1997
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 07/477,226
; FILING DATE: 08-FEB-1990
; APPLICATION NUMBER: 07/299,880
; FILING DATE: 19-JAN-1989
; APPLICATION NUMBER: 07/206,470
; FILING DATE: 14-JUN-1988
; APPLICATION NUMBER: 07/200,383
; FILING DATE: 31-MAY-1988
; ATTORNEY/AGENT INFORMATION:
; NAME: Murashige, Kate H
; REGISTRATION NUMBER: 29,959
; REFERENCE/DOCKET NUMBER: 219002025212
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 202-887-1500
; TELEFAX: 202-822-0168
; TELEX:
; INFORMATION FOR SEQ ID NO: 45:

; SEQUENCE CHARACTERISTICS:
; LENGTH: 134 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; FRAGMENT TYPE: internal
US-08-850-910A-45

Query Match 46.8%; Score 44; DB 1; Length 134;
Best Local Similarity 50.0%; Pred. No. 6.3;
Matches 9; Conservative 4; Mismatches 5; Indels 0; Gaps 0;

QY 1 QAOVNTIQGVSAALQTOQ 18
DB 44 QEQRNHLOKLSLQVEQ 61

RESULT 14
US-08-850-910A-48
; Sequence 48, Application US/08850910A
; Patent No. 5948761
; GENERAL INFORMATION:
; APPLICANT: SEILHAMER, J. J.
; APPLICANT: LEWICKI, J.
; APPLICANT: SCARBOROUGH, R. M.
; TITLE OF INVENTION: RECOMBINANT TECHNIQUES FOR
; TITLE OF INVENTION: PRODUCTION OF BRAIN NATRIUETIC PEPTIDE
; NUMBER OF SEQUENCES: 50
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: MORRISON & FOERSTER, LLP
; STREET: 2000 Pennsylvania Avenue, NW, Suite 5500
; CITY: Washington
; STATE: DC
; COUNTRY: USA
; ZIP: 20006-1888
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: Windows
; SOFTWARE: FastSeq for Windows Version 2.0b
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/850,910A
; FILING DATE: 05-MAY-1997
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 07/477,226
; FILING DATE: 08-FEB-1990
; APPLICATION NUMBER: 07/299,880
; FILING DATE: 19-JAN-1989
; APPLICATION NUMBER: 07/206,470
; FILING DATE: 14-JUN-1988
; APPLICATION NUMBER: 07/200,383
; FILING DATE: 31-MAY-1988
; ATTORNEY/AGENT INFORMATION:
; NAME: Murashige, Kate H
; REGISTRATION NUMBER: 29,959
; REFERENCE/DOCKET NUMBER: 219002025212
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 202-887-1500
; TELEFAX: 202-822-0168
; TELEX:
; INFORMATION FOR SEQ ID NO: 48:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 134 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
US-08-850-910A-48

Query Match 46.8%; Score 44; DB 1; Length 134;
Best Local Similarity 50.0%; Pred. No. 6.3;

Matches 9; Conservative 4; Mismatches 5; Indels 0; Gaps 0;

Qy 1 QAQVNTIQGVSAQTQQ 18
 Db 44 QQRNHLQGLSELOVEQ 61

RESULT 15
 US-09-508-435A-2
 ; Sequence 2, Application US/09508435A
 ; Patent No. 6828107
 ; GENERAL INFORMATION:
 ; APPLICANT: Shionogi & Co., Ltd.
 ; TITLE OF INVENTION: Immunoassay for BNP
 ; FILE REFERENCE: 2000-0259A/JJF/WMC/00177
 ; CURRENT APPLICATION NUMBER: US/09/508,435A
 ; CURRENT FILING DATE: 2000-03-13
 ; PRIOR APPLICATION NUMBER: JP 246684/1997
 ; PRIOR FILING DATE: 1997-09-11
 ; NUMBER OF SEQ ID NOS: 2
 ; SOFTWARE: Word (MS-DOS text)
 ; SEQ ID NO 2
 ; LENGTH: 134
 ; TYPE: PRT
 ; ORGANISM: human
 US-09-508-435A-2

Query Match 46.8%; Score 44; DB 2; Length 134;
 Best Local Similarity 50.0%; Pred. No. 6.3;
 Matches 9; Conservative 4; Mismatches 5; Indels 0; Gaps 0;

Qy 1 QAQVNTIQGVSAQTQQ 18
 Db 44 QQRNHLQGLSELOVEQ 61

Search completed: January 25, 2006, 19:18:15
 Job time : 17.6 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2006 Compugen Ltd.

OM protein - protein search, using sw model

Run on: January 25, 2006, 19:13:51 ; Search time 60 Seconds
(without alignments)
139.276 Million cell updates/sec

Title: US-10-797-821-9

Perfect score: 94

Sequence: 1 QAQVNTIQGVSAQTQQA 20

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1867569 seqs, 417829326 residues

Total number of hits satisfying chosen parameters: 1867569

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications AA Main:

- 1: /cgn2_6/prodata/1/pubpaa/US07_PUBCOMB.pap.*
- 2: /cgn2_6/prodata/1/pubpaa/US08_PUBCOMB.pap.*
- 3: /cgn2_6/prodata/1/pubpaa/US09_PUBCOMB.pap.*
- 4: /cgn2_6/prodata/1/pubpaa/US10A_PUBCOMB.pap.*
- 5: /cgn2_6/prodata/1/pubpaa/US10B_PUBCOMB.pap.*
- 6: /cgn2_6/prodata/1/pubpaa/US11_PUBCOMB.pap.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	94	100.0	20	4	US-10-383-930-9
2	94	100.0	20	5	US-10-797-821-9
3	94	100.0	431	4	US-10-383-930-29
4	94	100.0	431	4	US-10-383-930-30
5	94	100.0	431	4	US-10-383-930-33
6	94	100.0	431	5	US-10-797-821-29
7	94	100.0	431	5	US-10-797-821-30
8	94	100.0	431	5	US-10-797-821-33
9	94	100.0	432	4	US-10-383-930-31
10	94	100.0	432	4	US-10-383-930-32
11	94	100.0	432	5	US-10-797-821-31
12	94	100.0	432	5	US-10-797-821-32
13	75	79.8	20	4	US-10-383-930-8
14	75	79.8	20	5	US-10-797-821-8
15	65	69.1	20	4	US-10-383-930-2
16	65	69.1	20	5	US-10-797-821-2
17	53	56.4	218	4	US-10-437-963-199647
18	52	55.3	312	4	US-10-154-251-67
19	52	55.3	390	5	US-10-472-928-4652
20	52	55.3	399	5	US-10-617-320-3230
21	51	54.3	398	5	US-10-474-792-600
22	47	50.0	590	4	US-10-156-761-10807
23	47	50.0	1024	4	US-10-211-962-45
24	46	48.9	243	4	US-10-425-115-213685
25	45	47.9	198	4	US-10-424-599-164844
26	45	47.9	842	4	US-10-282-122A-65291
27	45	47.9	1024	4	US-10-211-962-43

28	45	47.9	1211	4	US-10-282-122A-69274	Sequence 69274, A
29	45	47.9	1379	6	US-11-097-143-261	Sequence 261, App
30	44	46.8	81	4	US-10-359-028-9	Sequence 9, Appli
31	44	46.8	108	4	US-10-419-059-1	Sequence 1, Appli
32	44	46.8	108	4	US-10-300-733-1	Sequence 1, Appli
33	44	46.8	108	4	US-10-359-051-1	Sequence 1, Appli
34	44	46.8	108	4	US-10-645-874-1	Sequence 1, Appli
35	44	46.8	108	5	US-10-714-078-1	Sequence 1, Appli
36	44	46.8	108	5	US-10-603-891-1	Sequence 1, Appli
37	44	46.8	108	5	US-10-827-919-1	Sequence 1, Appli
38	44	46.8	108	5	US-10-938-760-1	Sequence 1, Appli
39	44	46.8	108	5	US-10-952-275-1	Sequence 1, Appli
40	44	46.8	111	4	US-10-723-933-29	Sequence 29, Appl
41	44	46.8	134	3	US-09-902-517-45	Sequence 45, Appl
42	44	46.8	134	3	US-09-902-517-48	Sequence 48, Appl
43	44	46.8	134	3	US-09-508-435-2	Sequence 2, Appli
44	44	46.8	134	4	US-10-419-059-2	Sequence 2, Appli
45	44	46.8	134	4	US-10-402-021-45	Sequence 45, Appl

ALIGNMENTS

RESULT 1

US-10-383-930-9
; Sequence 9, Application US/10383930
; Publication No. US20040127400A1
; GENERAL INFORMATION:
; APPLICANT: Smith, Daniel J
; APPLICANT: Taubman, Martin A
; TITLE OF INVENTION: Immunogenicity of Glucan Binding Protein
; FILE REFERENCE: 25669-018
; CURRENT APPLICATION NUMBER: US/10/383,930
; CURRENT FILING DATE: 2003-03-07
; PRIOR APPLICATION NUMBER: 60/402,483
; PRIOR FILING DATE: 2002-08-08
; PRIOR APPLICATION NUMBER: 60/363,209
; PRIOR FILING DATE: 2002-03-07
; NUMBER OF SEQ ID NOS: 41
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 9
; LENGTH: 20
; TYPE: PRT
; ORGANISM: Streptococcus mutans
US-10-383-930-9

Query Match 100.0%; Score 94; DB 4; Length 20;
Best Local Similarity 100.0%; Pred. No. 7e-08; 0; Indels 0; Gaps 0;
Matches 20; Conservative 0; Mismatches 0;

QY 1 QAQVNTIQGVSAQTQQA 20
|||
DB 1 QAQVNTIQGVSAQTQQA 20
|||

RESULT 2

US-10-797-821-9
; Sequence 9, Application US/10797821
; Publication No. US20050031633A1
; GENERAL INFORMATION:
; APPLICANT: Smith, Daniel J.
; APPLICANT: Taubman, Martin A.
; TITLE OF INVENTION: Glucan Binding Protein and Glycosyltransferase Immunogens
; FILE REFERENCE: 25669-020
; CURRENT APPLICATION NUMBER: US/10/797,821
; CURRENT FILING DATE: 2004-03-09
; PRIOR APPLICATION NUMBER: 10/383,930
; PRIOR FILING DATE: 2003-03-07
; PRIOR APPLICATION NUMBER: 60/363,209
; PRIOR FILING DATE: 2002-03-07
; PRIOR APPLICATION NUMBER: 60/402,483
; PRIOR FILING DATE: 2002-08-08
; PRIOR APPLICATION NUMBER: 09/290,049

; PRIOR FILING DATE: 1999-04-12
; PRIOR APPLICATION NUMBER: 60/081,550
; PRIOR FILING DATE: 1998-04-13
; PRIOR APPLICATION NUMBER: 60/115,142
; PRIOR FILING DATE: 1999-01-08
; NUMBER OF SEQ ID NOS: 45
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 9
; LENGTH: 20
; TYPE: PRT
; ORGANISM: Artificial
; FEATURE:
; OTHER INFORMATION: GbpB peptide
US-10-797-821-9

Query Match 100.0%; Score 94; DB 5; Length 20;
Best Local Similarity 100.0%; Pred. No. 7e-08;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 QAVNTIQQVSALQTQQA 20
|||
Db 1 QAVNTIQQVSALQTQQA 20

RESULT 3
US-10-383-930-29
; Sequence 29, Application US/10383930
; Publication No. US20040127400A1
; GENERAL INFORMATION:
; APPLICANT: Smith, Daniel J
; APPLICANT: Taubman, Martin A
; TITLE OF INVENTION: Immunogenicity of Glucan Binding Protein
; FILE REFERENCE: 25669-018
; CURRENT APPLICATION NUMBER: US/10/383,930
; CURRENT FILING DATE: 2003-03-07
; PRIOR APPLICATION NUMBER: 60/402,483
; PRIOR FILING DATE: 2002-08-08
; PRIOR APPLICATION NUMBER: 60/363,209
; PRIOR FILING DATE: 2002-03-07
; NUMBER OF SEQ ID NOS: 41
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 29
; LENGTH: 431
; TYPE: PRT
; ORGANISM: Streptococcus mutans
US-10-383-930-29

Query Match 100.0%; Score 94; DB 4; Length 431;
Best Local Similarity 100.0%; Pred. No. 2.1e-06;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 QAVNTIQQVSALQTQQA 20
|||
Db 52 QAVNTIQQVSALQTQQA 71

RESULT 4
US-10-383-930-30
; Sequence 30, Application US/10383930
; Publication No. US20040127400A1
; GENERAL INFORMATION:
; APPLICANT: Smith, Daniel J
; APPLICANT: Taubman, Martin A
; TITLE OF INVENTION: Immunogenicity of Glucan Binding Protein
; FILE REFERENCE: 25669-018
; CURRENT APPLICATION NUMBER: US/10/383,930
; CURRENT FILING DATE: 2003-03-07
; PRIOR APPLICATION NUMBER: 60/402,483
; PRIOR FILING DATE: 2002-08-08
; PRIOR APPLICATION NUMBER: 60/363,209
; PRIOR FILING DATE: 2002-03-07
; NUMBER OF SEQ ID NOS: 41
; SOFTWARE: PatentIn version 3.2

; SEQ ID NO 30
; LENGTH: 431
; TYPE: PRT
; ORGANISM: Streptococcus mutans
US-10-383-930-30

Query Match 100.0%; Score 94; DB 4; Length 431;
Best Local Similarity 100.0%; Pred. No. 2.1e-06;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 QAVNTIQQVSALQTQQA 20
|||
Db 52 QAVNTIQQVSALQTQQA 71

RESULT 5
US-10-383-930-33
; Sequence 33, Application US/10383930
; Publication No. US20040127400A1
; GENERAL INFORMATION:
; APPLICANT: Smith, Daniel J
; APPLICANT: Taubman, Martin A
; TITLE OF INVENTION: Immunogenicity of Glucan Binding Protein
; FILE REFERENCE: 25669-018
; CURRENT APPLICATION NUMBER: US/10/383,930
; CURRENT FILING DATE: 2003-03-07
; PRIOR APPLICATION NUMBER: 60/402,483
; PRIOR FILING DATE: 2002-08-08
; PRIOR APPLICATION NUMBER: 60/363,209
; PRIOR FILING DATE: 2002-03-07
; NUMBER OF SEQ ID NOS: 41
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 33
; LENGTH: 431
; TYPE: PRT
; ORGANISM: Streptococcus mutans
US-10-383-930-33

Query Match 100.0%; Score 94; DB 4; Length 431;
Best Local Similarity 100.0%; Pred. No. 2.1e-06;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 QAVNTIQQVSALQTQQA 20
|||
Db 52 QAVNTIQQVSALQTQQA 71

RESULT 6
US-10-797-821-29
; Sequence 29, Application US/10797821
; Publication No. US20050031633A1
; GENERAL INFORMATION:
; APPLICANT: Smith, Daniel J.
; APPLICANT: Taubman, Martin A.
; TITLE OF INVENTION: Glucan Binding Protein and Glycosyltransferase Immunogens
; FILE REFERENCE: 25669-020
; CURRENT APPLICATION NUMBER: US/10/797,821
; CURRENT FILING DATE: 2004-03-09
; PRIOR APPLICATION NUMBER: 10/383,930
; PRIOR FILING DATE: 2003-03-07
; PRIOR APPLICATION NUMBER: 60/363,209
; PRIOR FILING DATE: 2002-03-07
; PRIOR APPLICATION NUMBER: 60/402,483
; PRIOR FILING DATE: 2002-08-08
; PRIOR APPLICATION NUMBER: 09/290,049
; PRIOR FILING DATE: 1999-04-12
; PRIOR APPLICATION NUMBER: 60/081,550
; PRIOR FILING DATE: 1998-04-13
; PRIOR APPLICATION NUMBER: 60/115,142
; PRIOR FILING DATE: 1999-01-08
; NUMBER OF SEQ ID NOS: 45
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 29

```

; LENGTH: 431
; TYPE: PRT
; ORGANISM: Streptococcus mutans
US-10-797-821-29

Query Match
Best Local Similarity 100.0%; Score 94; DB 5; Length 431;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 QAVNTIQQVSALQTQQA 20
Db 52 QAVNTIQQVSALQTQQA 71

RESULT 7
US-10-797-821-30
; Sequence 30, Application US/10797821
; Publication No. US20050031633A1
; GENERAL INFORMATION:
; APPLICANT: Smith, Daniel J.
; APPLICANT: Taubman, Martin A.
; TITLE OF INVENTION: Glucan Binding Protein and Glycosyltransferase Immunogens
; FILE REFERENCE: 25669-020
; CURRENT APPLICATION NUMBER: US/10/797,821
; CURRENT FILING DATE: 2004-03-09
; PRIOR APPLICATION NUMBER: 10/383,930
; PRIOR FILING DATE: 2003-03-07
; PRIOR APPLICATION NUMBER: 60/363,209
; PRIOR FILING DATE: 2002-03-07
; PRIOR APPLICATION NUMBER: 60/402,483
; PRIOR FILING DATE: 2002-08-08
; PRIOR APPLICATION NUMBER: 09/290,049
; PRIOR FILING DATE: 1999-04-12
; PRIOR APPLICATION NUMBER: 60/081,550
; PRIOR FILING DATE: 1998-04-13
; PRIOR APPLICATION NUMBER: 60/115,142
; PRIOR FILING DATE: 1999-01-08
; NUMBER OF SEQ ID NOS: 45
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 30
; LENGTH: 431
; TYPE: PRT
; ORGANISM: Streptococcus mutans
US-10-797-821-30

Query Match
Best Local Similarity 100.0%; Score 94; DB 5; Length 431;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 QAVNTIQQVSALQTQQA 20
Db 52 QAVNTIQQVSALQTQQA 71

RESULT 8
US-10-797-821-33
; Sequence 33, Application US/10797821
; Publication No. US20050031633A1
; GENERAL INFORMATION:
; APPLICANT: Smith, Daniel J.
; APPLICANT: Taubman, Martin A.
; TITLE OF INVENTION: Glucan Binding Protein and Glycosyltransferase Immunogens
; FILE REFERENCE: 25669-020
; CURRENT APPLICATION NUMBER: US/10/797,821
; CURRENT FILING DATE: 2004-03-09
; PRIOR APPLICATION NUMBER: 10/383,930
; PRIOR FILING DATE: 2003-03-07
; PRIOR APPLICATION NUMBER: 60/363,209
; PRIOR FILING DATE: 2002-03-07
; PRIOR APPLICATION NUMBER: 60/402,483
; PRIOR FILING DATE: 2002-08-08
; PRIOR APPLICATION NUMBER: 09/290,049
; PRIOR FILING DATE: 1999-04-12

```

```

; PRIOR APPLICATION NUMBER: 60/081,550
; PRIOR FILING DATE: 1998-04-13
; PRIOR APPLICATION NUMBER: 60/115,142
; PRIOR FILING DATE: 1999-01-08
; NUMBER OF SEQ ID NOS: 45
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 33
; LENGTH: 431
; TYPE: PRT
; ORGANISM: Streptococcus mutans
US-10-797-821-33

Query Match
Best Local Similarity 100.0%; Score 94; DB 5; Length 431;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 QAVNTIQQVSALQTQQA 20
Db 52 QAVNTIQQVSALQTQQA 71

RESULT 9
US-10-383-930-31
; Sequence 31, Application US/10383930
; Publication No. US20040127400A1
; GENERAL INFORMATION:
; APPLICANT: Smith, Daniel J.
; APPLICANT: Taubman, Martin A.
; TITLE OF INVENTION: Immunogenicity of Glucan Binding Protein
; FILE REFERENCE: 25669-018
; CURRENT APPLICATION NUMBER: US/10/383,930
; CURRENT FILING DATE: 2003-03-07
; PRIOR APPLICATION NUMBER: 60/402,483
; PRIOR FILING DATE: 2002-08-08
; PRIOR APPLICATION NUMBER: 60/363,209
; PRIOR FILING DATE: 2002-03-07
; NUMBER OF SEQ ID NOS: 41
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 31
; LENGTH: 432
; TYPE: PRT
; ORGANISM: Streptococcus mutans
US-10-383-930-31

Query Match
Best Local Similarity 100.0%; Score 94; DB 4; Length 432;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 QAVNTIQQVSALQTQQA 20
Db 52 QAVNTIQQVSALQTQQA 71

RESULT 10
US-10-383-930-32
; Sequence 32, Application US/10383930
; Publication No. US20040127400A1
; GENERAL INFORMATION:
; APPLICANT: Smith, Daniel J.
; APPLICANT: Taubman, Martin A.
; TITLE OF INVENTION: Immunogenicity of Glucan Binding Protein
; FILE REFERENCE: 25669-018
; CURRENT APPLICATION NUMBER: US/10/383,930
; CURRENT FILING DATE: 2003-03-07
; PRIOR APPLICATION NUMBER: 60/402,483
; PRIOR FILING DATE: 2002-08-08
; PRIOR APPLICATION NUMBER: 60/363,209
; PRIOR FILING DATE: 2002-03-07
; NUMBER OF SEQ ID NOS: 41
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 32
; LENGTH: 432
; TYPE: PRT

```

```

; ORGANISM: Streptococcus mutans
US-10-383-930-32

Query Match      100.0%; Score 94; DB 4; Length 432;
Best Local Similarity 100.0%; Pred. No. 2.1e-06;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 QAQVNTIQGVSAALQTQAAE 20
   |||||
Db 52 QAQVNTIQGVSAALQTQAAE 71

RESULT 11
US-10-797-821-31
; Sequence 31, Application US/10797821
; Publication No. US20050031633A1
; GENERAL INFORMATION:
; APPLICANT: Smith, Daniel J.
; APPLICANT: Taubman, Martin A.
; TITLE OF INVENTION: Glucan Binding Protein and Glycosyltransferase Immunogens
; FILE REFERENCE: 25669-020
; CURRENT APPLICATION NUMBER: US/10797,821
; CURRENT FILING DATE: 2004-03-09
; PRIOR APPLICATION NUMBER: 10/383,930
; PRIOR FILING DATE: 2003-03-07
; PRIOR APPLICATION NUMBER: 60/363,209
; PRIOR FILING DATE: 2002-03-07
; PRIOR APPLICATION NUMBER: 60/402,483
; PRIOR FILING DATE: 2002-08-08
; PRIOR APPLICATION NUMBER: 09/290,049
; PRIOR FILING DATE: 1999-04-12
; PRIOR APPLICATION NUMBER: 60/081,550
; PRIOR FILING DATE: 1998-04-13
; PRIOR APPLICATION NUMBER: 60/115,142
; PRIOR FILING DATE: 1999-01-08
; NUMBER OF SEQ ID NOS: 45
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 31
; LENGTH: 432
; TYPE: PRT
; ORGANISM: Streptococcus mutans
US-10-797-821-31

Query Match      100.0%; Score 94; DB 5; Length 432;
Best Local Similarity 100.0%; Pred. No. 2.1e-06;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 QAQVNTIQGVSAALQTQAAE 20
   |||||
Db 52 QAQVNTIQGVSAALQTQAAE 71

RESULT 12
US-10-797-821-32
; Sequence 32, Application US/10797821
; Publication No. US20050031633A1
; GENERAL INFORMATION:
; APPLICANT: Smith, Daniel J.
; APPLICANT: Taubman, Martin A.
; TITLE OF INVENTION: Glucan Binding Protein and Glycosyltransferase Immunogens
; FILE REFERENCE: 25669-020
; CURRENT APPLICATION NUMBER: US/10797,821
; CURRENT FILING DATE: 2004-03-09
; PRIOR APPLICATION NUMBER: 10/383,930
; PRIOR FILING DATE: 2003-03-07
; PRIOR APPLICATION NUMBER: 60/363,209
; PRIOR FILING DATE: 2002-03-07
; PRIOR APPLICATION NUMBER: 60/402,483
; PRIOR FILING DATE: 2002-08-08
; PRIOR APPLICATION NUMBER: 09/290,049
; PRIOR FILING DATE: 1999-04-12
; PRIOR APPLICATION NUMBER: 60/081,550
; PRIOR FILING DATE: 1998-04-13

```

```

; PRIOR APPLICATION NUMBER: 60/115,142
; PRIOR FILING DATE: 1999-01-08
; NUMBER OF SEQ ID NOS: 45
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 32
; LENGTH: 432
; TYPE: PRT
; ORGANISM: Streptococcus mutans
US-10-797-821-32

Query Match      100.0%; Score 94; DB 5; Length 432;
Best Local Similarity 100.0%; Pred. No. 2.1e-06;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 QAQVNTIQGVSAALQTQAAE 20
   |||||
Db 52 QAQVNTIQGVSAALQTQAAE 71

RESULT 13
US-10-383-930-8
; Sequence 8, Application US/10383930
; Publication No. US20040127400A1
; GENERAL INFORMATION:
; APPLICANT: Smith, Daniel J.
; APPLICANT: Taubman, Martin A.
; TITLE OF INVENTION: Immunogenicity of Glucan Binding Protein
; FILE REFERENCE: 25669-018
; CURRENT APPLICATION NUMBER: US/10/383,930
; CURRENT FILING DATE: 2003-03-07
; PRIOR APPLICATION NUMBER: 60/402,483
; PRIOR FILING DATE: 2002-08-08
; PRIOR APPLICATION NUMBER: 60/363,209
; PRIOR FILING DATE: 2002-03-07
; NUMBER OF SEQ ID NOS: 41
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 8
; LENGTH: 20
; TYPE: PRT
; ORGANISM: Streptococcus mutans
US-10-383-930-8

Query Match      79.8%; Score 75; DB 4; Length 20;
Best Local Similarity 100.0%; Pred. No. 7e-05;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 QAQVNTIQGVSAALQT 16
   |||||
Db 5 QAQVNTIQGVSAALQT 20

RESULT 14
US-10-797-821-8
; Sequence 8, Application US/10797821
; Publication No. US20050031633A1
; GENERAL INFORMATION:
; APPLICANT: Smith, Daniel J.
; APPLICANT: Taubman, Martin A.
; TITLE OF INVENTION: Glucan Binding Protein and Glycosyltransferase Immunogens
; FILE REFERENCE: 25669-020
; CURRENT APPLICATION NUMBER: US/10797,821
; CURRENT FILING DATE: 2004-03-09
; PRIOR APPLICATION NUMBER: 10/383,930
; PRIOR FILING DATE: 2003-03-07
; PRIOR APPLICATION NUMBER: 60/363,209
; PRIOR FILING DATE: 2002-03-07
; PRIOR APPLICATION NUMBER: 60/402,483
; PRIOR FILING DATE: 2002-08-08
; PRIOR APPLICATION NUMBER: 09/290,049
; PRIOR FILING DATE: 1999-04-12
; PRIOR APPLICATION NUMBER: 60/081,550
; PRIOR FILING DATE: 1998-04-13
; PRIOR APPLICATION NUMBER: 60/115,142

```

; PRIOR FILING DATE: 1999-01-08
 ; NUMBER OF SEQ ID NOS: 45
 ; SOFTWARE: PatentIn version 3.2
 ; SEQ ID NO 8
 ; LENGTH: 20
 ; TYPE: PRT
 ; ORGANISM: Artificial
 ; FEATURE:
 ; OTHER INFORMATION: GbpB peptide
 US-10-797-821-8

Query Match 79.8%; Score 75; DB 5; Length 20;
 Best Local Similarity 100.0%; Pred. No. 7e-05;
 Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 QAVNTIQQVSALQT 16
 |||||
 Db 5 QAVNTIQQVSALQT 20

RESULT 15
 US-10-383-930-2
 ; Sequence 2, Application US/10383930
 ; Publication No. US20040127400A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Smith, Daniel J
 ; APPLICANT: Taubman, Martin A
 ; TITLE OF INVENTION: Immunogenicity of Glucan Binding Protein
 ; FILE REFERENCE: 25669-018
 ; CURRENT APPLICATION NUMBER: US/10/383,930
 ; CURRENT FILING DATE: 2003-03-07
 ; PRIOR APPLICATION NUMBER: 60/402,483
 ; PRIOR FILING DATE: 2002-08-08
 ; PRIOR APPLICATION NUMBER: 60/363,209
 ; PRIOR FILING DATE: 2002-03-07
 ; NUMBER OF SEQ ID NOS: 41
 ; SOFTWARE: PatentIn version 3.2
 ; SEQ ID NO 2
 ; LENGTH: 20
 ; TYPE: PRT
 ; ORGANISM: Streptococcus mutans
 US-10-383-930-2

Query Match 69.1%; Score 65; DB 4; Length 20;
 Best Local Similarity 100.0%; Pred. No. 0.0027;
 Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 6 TIQGVVSALQTQQA 19
 |||||
 Db 6 TIQGVVSALQTQQA 19

Search completed: January 25, 2006, 20:10:04
 Job time : 60 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2006 Compugen Ltd.

OM protein - protein search, using sw model

Run on: January 25, 2006, 19:15:26 ; Search time 5.5 Seconds
(without alignments)
39.378 Million cell updates/sec

Title: US-10-797-821-9
Perfect score: 94
Sequence: 1 QAOVNTIQVSAQTQQA 20

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 75621 seqs, 10829074 residues

Total number of hits satisfying chosen parameters: 75621

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published Applications AA New.*
1: /cgn2_6/prodata/2/pubpaa/US08_NEW_PUB.pep.*
2: /cgn2_6/prodata/2/pubpaa/US06_NEW_PUB.pep.*
3: /cgn2_6/prodata/2/pubpaa/US07_NEW_PUB.pep.*
4: /cgn2_6/prodata/2/pubpaa/PCT_NEW_PUB.pep.*
5: /cgn2_6/prodata/2/pubpaa/US09_NEW_PUB.pep.*
6: /cgn2_6/prodata/2/pubpaa/US10_NEW_PUB.pep.*
7: /cgn2_6/prodata/2/pubpaa/US11_NEW_PUB.pep.*
8: /cgn2_6/prodata/2/pubpaa/US60_NEW_PUB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	94	100.0	431	7	US-11-052-554A-210
2	51	54.3	398	7	US-11-052-554A-252
3	46	48.9	768	7	US-11-147-109-10
4	45	47.9	642	6	US-10-467-657-5304
5	44	46.8	76	6	US-10-510-880-6
6	44	46.8	76	6	US-10-977-334-2
7	44	46.8	108	6	US-10-510-880-5
8	44	46.8	108	6	US-10-875-800-1
9	44	46.8	108	6	US-10-299-977-1
10	44	46.8	134	6	US-10-510-880-1
11	44	46.8	134	6	US-10-875-800-2
12	44	46.8	134	6	US-11-043-590-12
13	44	46.8	162	7	US-11-043-590-13
14	44	46.8	3674	7	US-11-000-463-454
15	42	44.7	1392	7	US-11-069-834-58
16	42	44.7	2000	7	US-11-069-834-56
17	41	43.6	1995	7	US-11-069-834-60
18	41	43.6	2101	6	US-10-857-780-23
19	40	42.6	202	6	US-10-454-437-62
20	40	42.6	392	6	US-10-467-657-5980
21	40	42.6	472	6	US-10-508-263-62
22	40	42.6	472	6	US-10-508-263-64
23	40	42.6	488	6	US-10-508-263-12
24	40	42.6	490	6	US-10-508-263-10
25	40	42.6	544	6	US-10-821-234-889

26	40	42.6	989	6	US-10-821-234-975	Sequence 975, App
27	39.5	42.0	1122	6	US-10-467-657-6112	Sequence 6112, App
28	39	41.5	205	6	US-10-467-657-4336	Sequence 4336, App
29	39	41.5	431	6	US-10-821-234-1285	Sequence 1285, App
30	39	41.5	451	7	US-11-052-554A-73	Sequence 73, Appl
31	39	41.5	455	6	US-10-508-263-66	Sequence 66, Appl
32	39	41.5	706	6	US-10-485-517-146	Sequence 146, Appl
33	39	41.5	908	7	US-11-077-550-64	Sequence 64, Appl
34	39	41.5	914	7	US-11-077-550-122	Sequence 122, App
35	39	41.5	944	7	US-11-077-550-118	Sequence 118, App
36	39	41.5	950	7	US-10-821-234-1648	Sequence 1648, App
37	39	41.5	1304	6	US-11-091-668-4	Sequence 4, Appli
38	38.5	41.0	1313	7	US-11-113-581-88	Sequence 88, Appl
39	38	40.4	175	7	US-11-113-581-97	Sequence 97, Appl
40	38	40.4	184	6	US-10-508-263-16	Sequence 16, Appl
41	38	40.4	228	6	US-10-467-657-568	Sequence 568, App
42	38	40.4	228	6	US-10-467-657-4838	Sequence 4838, App
43	38	40.4	264	7	US-11-188-743-24	Sequence 24, Appl
44	38	40.4	264	7	US-11-188-743-25	Sequence 25, Appl
45	38	40.4				

ALIGNMENTS

RESULT 1
US-11-052-554A-210
; Sequence 210, Application US/11052554A
; Publication No. US20050288866A1
; GENERAL INFORMATION:
; APPLICANT: Sachdeva, et al.
; TITLE OF INVENTION: COMPUTATIONAL METHOD FOR IDENTIFYING ADHESIN AND ADHESIN-LIKE
; TITLE OF INVENTION: PROTEINS OF THERAPEUTIC POTENTIAL
; FILE REFERENCE: 30853/40359A
; CURRENT APPLICATION NUMBER: US/11/052,554A
; CURRENT FILING DATE: 2005-02-07
; PRIOR APPLICATION NUMBER: US 60/589,227
; PRIOR FILING DATE: 2004-07-20
; PRIOR APPLICATION NUMBER: IN 173/DEL/2004
; PRIOR FILING DATE: 2004-02-06
; NUMBER OF SEQ ID NOS: 763
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 210
; LENGTH: 431
; TYPE: PRT
; ORGANISM: Streptococcus mutans UA159
US-11-052-554A-210

Query Match 100.0%; Score 94; DB 7; Length 431;
Best Local Similarity 100.0%; Pred. No. 3.3e-08;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 QAOVNTIQVSAQTQQA 20
Db 52 QAOVNTIQVSAQTQQA 71

RESULT 2
US-11-052-554A-252
; Sequence 252, Application US/11052554A
; Publication No. US20050288866A1
; GENERAL INFORMATION:
; APPLICANT: Sachdeva, et al.
; TITLE OF INVENTION: COMPUTATIONAL METHOD FOR IDENTIFYING ADHESIN AND ADHESIN-LIKE
; TITLE OF INVENTION: PROTEINS OF THERAPEUTIC POTENTIAL
; FILE REFERENCE: 30853/40359A
; CURRENT APPLICATION NUMBER: US/11/052,554A
; CURRENT FILING DATE: 2005-02-07
; PRIOR APPLICATION NUMBER: US 60/589,227
; PRIOR FILING DATE: 2004-07-20
; PRIOR APPLICATION NUMBER: IN 173/DEL/2004
; PRIOR FILING DATE: 2004-02-06
; NUMBER OF SEQ ID NOS: 763

```
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 252
; LENGTH: 398
; TYPE: PRT
; ORGANISM: Streptococcus pyogenes MGAS8232
US-11-052-554A-252

Query Match      54.3%; Score 51; DB 7; Length 398;
Best Local Similarity 55.6%; Pred. No. 0.32;
Matches 10; Conservative 5; Mismatches 3; Indels 0; Gaps 0;

QY 1 QAQVNTIQGVSAALQTQAE 18
Db 49 QNQVSALQAQVSSLSQSEQ 66

RESULT 3
US-11-147-109-10
; Sequence 10, Application US/11147109
; Publication No. US20050278805A1
; GENERAL INFORMATION:
; APPLICANT: Shen, Bo
; APPLICANT: Tarczynski, Mitchell C.
; TITLE OF INVENTION: Method of Increasing Seed Oil Content in
; FILE REFERENCE: 1409P
; CURRENT APPLICATION NUMBER: US/11/147,109
; CURRENT FILING DATE: 2005-06-07
; NUMBER OF SEQ ID NOS: 10
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 10
; LENGTH: 768
; TYPE: PRT
; ORGANISM: Phalaenopsis sp.
US-11-147-109-10

Query Match      48.9%; Score 46; DB 7; Length 768;
Best Local Similarity 52.6%; Pred. No. 4.4;
Matches 10; Conservative 3; Mismatches 6; Indels 0; Gaps 0;

QY 2 AQVNTIQGVSAALQTQAE 20
Db 26 AQINILEGQQLPQHQLAE 44

RESULT 4
US-10-467-657-5304
; Sequence 5304, Application US/10467657
; Publication No. US20050260581A1
; GENERAL INFORMATION:
; APPLICANT: CHIRON SPA
; APPLICANT: FONTANA Maria Rita
; APPLICANT: PIZZA Mariagrazia
; APPLICANT: MASIGNANI Vega
; APPLICANT: MONACI Elisabetta
; TITLE OF INVENTION: GONOCOCCAL PROTEINS AND NUCLEIC ACIDS
; FILE REFERENCE:
; CURRENT APPLICATION NUMBER: US/10/467,657
; CURRENT FILING DATE: 2003-08-11
; PRIOR APPLICATION NUMBER: GB-0103424.8
; PRIOR FILING DATE: 2001-02-12
; NUMBER OF SEQ ID NOS: 9218
; SOFTWARE: SeqWin99, version 1.04
; SEQ ID NO 5304
; LENGTH: 642
; TYPE: PRT
; ORGANISM: Neisseria gonorrhoeae
US-10-467-657-5304

Query Match      47.9%; Score 45; DB 6; Length 642;
Best Local Similarity 35.0%; Pred. No. 5.2;
Matches 7; Conservative 9; Mismatches 4; Indels 0; Gaps 0;

; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 252
; LENGTH: 592
; TYPE: PRT
; ORGANISM: Streptococcus pyogenes MGAS8232
US-10-052-554A-252

Query Match      54.3%; Score 51; DB 7; Length 398;
Best Local Similarity 55.6%; Pred. No. 0.32;
Matches 10; Conservative 5; Mismatches 3; Indels 0; Gaps 0;

QY 1 QAQVNTIQGVSAALQTQAE 20
Db 573 QRELDALPDEIAALETEQAE 592

RESULT 5
US-10-510-880-6
; Sequence 6, Application US/10510880
; Publication No. US20050244902A1
; GENERAL INFORMATION:
; APPLICANT: Rehfeld, Jens P.
; APPLICANT: Goetze, Jens Peter
; APPLICANT: Rigthospitalet
; TITLE OF INVENTION: Methods for determining levels of human
; TITLE OF INVENTION: B-type natriuretic peptide precursors
; FILE REFERENCE: 271212000200
; CURRENT APPLICATION NUMBER: US/10/510,880
; CURRENT FILING DATE: 2004-10-08
; PRIOR APPLICATION NUMBER: PCT/DK03/00250
; PRIOR FILING DATE: 2003-04-11
; PRIOR APPLICATION NUMBER: PSI692
; PRIOR FILING DATE: 2002-04-11
; NUMBER OF SEQ ID NOS: 7
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 6
; LENGTH: 76
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-510-880-6

Query Match      46.8%; Score 44; DB 6; Length 76;
Best Local Similarity 50.0%; Pred. No. 0.64;
Matches 9; Conservative 4; Mismatches 5; Indels 0; Gaps 0;

QY 1 QAQVNTIQGVSAALQTQ 18
Db 18 QEQRNHLQGLSELQVEQ 35

RESULT 6
US-10-977-334-2
; Sequence 2, Application US/10977334
; Publication No. US20050244904A1
; GENERAL INFORMATION:
; APPLICANT: NG, LEONG
; TITLE OF INVENTION: DIAGNOSTICS BASED ON SIGNAL PEPTIDE DETECTION
; FILE REFERENCE: ISA-016.01
; CURRENT APPLICATION NUMBER: US/10/977,334
; CURRENT FILING DATE: 2004-10-29
; PRIOR APPLICATION NUMBER: 60/542,647
; PRIOR FILING DATE: 2004-02-06
; PRIOR APPLICATION NUMBER: GB 0325279.8
; PRIOR FILING DATE: 2003-10-29
; NUMBER OF SEQ ID NOS: 8
; SOFTWARE: PatentIn Ver. 3.3
; SEQ ID NO 2
; LENGTH: 76
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-977-334-2

Query Match      46.8%; Score 44; DB 6; Length 76;
Best Local Similarity 50.0%; Pred. No. 0.64;
Matches 9; Conservative 4; Mismatches 5; Indels 0; Gaps 0;

QY 1 QAQVNTIQGVSAALQTQ 18
Db 18 QEQRNHLQGLSELQVEQ 35

RESULT 7
US-10-510-880-5
; Sequence 5, Application US/10510880
```

Publication No. US20050244902A1
; GENERAL INFORMATION:
; APPLICANT: Rehfeld, Jens F.
; APPLICANT: Goetze, Jens Peter
; APPLICANT: Rigshospitalet
; TITLE OF INVENTION: Methods for determining levels of human
; FILE REFERENCE: 271212000200
; CURRENT APPLICATION NUMBER: US/10/510,880
; CURRENT FILING DATE: 2004-10-08
; PRIOR APPLICATION NUMBER: PCT/DK03/00250
; PRIOR FILING DATE: 2003-04-11
; PRIOR APPLICATION NUMBER: PSI692
; PRIOR FILING DATE: 2002-04-11
; NUMBER OF SEQ ID NOS: 7
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 5
; LENGTH: 108
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-510-880-5

Query Match 46.8%; Score 44; DB 6; Length 108;
Best Local Similarity 50.0%; Pred. No. 0.97;
Matches 9; Conservative 4; Mismatches 5; Indels 0; Gaps 0;

Qy 1 QAQVNTIQGVSAQTQQ 18
Db 18 QEQRNHLQGLSELQVEQ 35

RESULT 8
US-10-875-800-1
; Sequence 1, Application US/10875800
; Publication No. US20050255484A1
; GENERAL INFORMATION:
; APPLICANT: VALKIRS, GUNARS
; APPLICANT: DAHLEN, JEFF
; APPLICANT: KIRCHICK, HOWARD
; APPLICANT: BUECHLER, KEN
; TITLE OF INVENTION: DIAGNOSTIC MARKERS OF STROKE AND CEREBRAL INJURY
; FILE REFERENCE: 071949-5408
; CURRENT APPLICATION NUMBER: US/10/875,800
; CURRENT FILING DATE: 2004-06-23
; PRIOR APPLICATION NUMBER: US/10/714,078
; PRIOR FILING DATE: 2003-11-14
; PRIOR APPLICATION NUMBER: 10/371,149
; PRIOR FILING DATE: 2003-02-20
; PRIOR APPLICATION NUMBER: 10/225,082
; PRIOR FILING DATE: 2002-08-20
; PRIOR APPLICATION NUMBER: PCT/US02/26604
; PRIOR FILING DATE: 2002-08-20
; PRIOR APPLICATION NUMBER: 60/313,775
; PRIOR FILING DATE: 2001-08-20
; PRIOR APPLICATION NUMBER: 60/334,964
; PRIOR FILING DATE: 2001-11-30
; PRIOR APPLICATION NUMBER: 60/346,485
; PRIOR FILING DATE: 2002-01-02
; NUMBER OF SEQ ID NOS: 7
; SOFTWARE: PatentIn Ver. 3.2
; SEQ ID NO 1
; LENGTH: 108
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-875-800-1

Query Match 46.8%; Score 44; DB 6; Length 108;
Best Local Similarity 50.0%; Pred. No. 0.97;
Matches 9; Conservative 4; Mismatches 5; Indels 0; Gaps 0;

Qy 1 QAQVNTIQGVSAQTQQ 18
Db 18 QEQRNHLQGLSELQVEQ 35

Db 18 QEQRNHLQGLSELQVEQ 35
RESULT 9
US-10-299-977-1
; Sequence 1, Application US/10299977
; Publication No. US20050287613A1
; GENERAL INFORMATION:
; APPLICANT: Jackowski, George
; APPLICANT: Kupchak, Peter
; APPLICANT: Stanton, Eric
; APPLICANT: Davey, Michelle
; TITLE OF INVENTION: Polyclonal-Polyclonal ELISA Assay For Detecting N-Terminus Problem
; FILE REFERENCE: 2132.127
; CURRENT APPLICATION NUMBER: US/10/299,977
; CURRENT FILING DATE: 2002-11-18
; NUMBER OF SEQ ID NOS: 1
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 1
; LENGTH: 108
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-299-977-1

Query Match 46.8%; Score 44; DB 6; Length 108;
Best Local Similarity 50.0%; Pred. No. 0.97;
Matches 9; Conservative 4; Mismatches 5; Indels 0; Gaps 0;

Qy 1 QAQVNTIQGVSAQTQQ 18
Db 18 QEQRNHLQGLSELQVEQ 35

RESULT 10
US-10-510-880-1
; Sequence 1, Application US/10510880
; Publication No. US20050244902A1
; GENERAL INFORMATION:
; APPLICANT: Rehfeld, Jens F.
; APPLICANT: Rigshospitalet
; TITLE OF INVENTION: Methods for determining levels of human
; FILE REFERENCE: 271212000200
; CURRENT APPLICATION NUMBER: US/10/510,880
; CURRENT FILING DATE: 2004-10-08
; PRIOR APPLICATION NUMBER: PCT/DK03/00250
; PRIOR FILING DATE: 2003-04-11
; PRIOR APPLICATION NUMBER: PSI692
; PRIOR FILING DATE: 2002-04-11
; NUMBER OF SEQ ID NOS: 7
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1
; LENGTH: 134
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-510-880-1

Query Match 46.8%; Score 44; DB 6; Length 134;
Best Local Similarity 50.0%; Pred. No. 1.2;
Matches 9; Conservative 4; Mismatches 5; Indels 0; Gaps 0;

Qy 1 QAQVNTIQGVSAQTQQ 18
Db 44 QEQRNHLQGLSELQVEQ 61

RESULT 11
US-10-875-800-2
; Sequence 2, Application US/10875800
; Publication No. US20050255484A1
; GENERAL INFORMATION:
; APPLICANT: VALKIRS, GUNARS

; APPLICANT: DAHLEN, JEFF
 ; APPLICANT: KIRCHICK, HOWARD
 ; APPLICANT: BUECHLER, KEN
 ; TITLE OF INVENTION: DIAGNOSTIC MARKERS OF STROKE AND CEREBRAL INJURY AND
 ; FILE REFERENCE: 071949-5408
 ; CURRENT APPLICATION NUMBER: US/10/875,800
 ; PRIOR FILING DATE: 2004-06-23
 ; PRIOR FILING DATE: 2003-11-14
 ; PRIOR FILING DATE: 2003-02-20
 ; PRIOR FILING DATE: 2003-02-20
 ; PRIOR FILING DATE: 2002-08-20
 ; PRIOR FILING DATE: 2002-08-20
 ; PRIOR FILING DATE: 2002-08-20
 ; PRIOR FILING DATE: 2001-08-20
 ; PRIOR FILING DATE: 2001-08-20
 ; PRIOR FILING DATE: 2001-11-30
 ; PRIOR FILING DATE: 2001-11-30
 ; PRIOR FILING DATE: 2002-01-02
 ; NUMBER OF SEQ ID NOS: 7
 ; SOFTWARE: PatentIn Ver. 3.2
 ; SEQ ID NO 2
 ; LENGTH: 134
 ; TYPE: PRT
 ; ORGANISM: Homo sapiens
 US-10-875-800-2

Query Match 46.8%; Score 44; DB 6; Length 134;
 Best Local Similarity 50.0%; Pred. No. 1.2;
 Matches 9; Conservative 4; Mismatches 5; Indels 0; Gaps 0;

QY 1 QAQVNTIQGVSAALQTQQ 18
 DB 44 QQRNHLQKLSLQVEQ 61

RESULT 12
 ; Sequence 12, Application US/11043590
 ; Publication No. US20050277156A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Compugen Ltd
 ; TITLE OF INVENTION: Novel Brain Natriuretic Peptide Variants and Methods of use there
 ; FILE REFERENCE: 1847.1011
 ; CURRENT APPLICATION NUMBER: US/11/043,590
 ; CURRENT FILING DATE: 2005-01-27
 ; NUMBER OF SEQ ID NOS: 47
 ; SEQ ID NO 12
 ; LENGTH: 134
 ; TYPE: PRT
 ; ORGANISM: Homo sapiens
 US-11-043-590-12

Query Match 46.8%; Score 44; DB 7; Length 134;
 Best Local Similarity 50.0%; Pred. No. 1.2;
 Matches 9; Conservative 4; Mismatches 5; Indels 0; Gaps 0;

QY 1 QAQVNTIQGVSAALQTQQ 18
 DB 44 QQRNHLQKLSLQVEQ 61

RESULT 13
 ; Sequence 13, Application US/11043590
 ; Publication No. US20050277156A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Compugen Ltd
 ; TITLE OF INVENTION: Novel Brain Natriuretic Peptide Variants and Methods of use there
 ; FILE REFERENCE: 1847.1011

; CURRENT APPLICATION NUMBER: US/11/043,590
 ; CURRENT FILING DATE: 2005-01-27
 ; NUMBER OF SEQ ID NOS: 47
 ; SEQ ID NO 13
 ; LENGTH: 162
 ; TYPE: PRT
 ; ORGANISM: Homo sapiens
 US-11-043-590-13

Query Match 46.8%; Score 44; DB 7; Length 162;
 Best Local Similarity 50.0%; Pred. No. 1.5;
 Matches 9; Conservative 4; Mismatches 5; Indels 0; Gaps 0;

QY 1 QAQVNTIQGVSAALQTQQ 18
 DB 44 QQRNHLQKLSLQVEQ 61

RESULT 14
 ; Sequence 454, Application US/110000463
 ; Publication No. US20050266423A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Tang, Y Tom
 ; APPLICANT: Liu, Chenghua
 ; APPLICANT: Asundi, Vinod
 ; APPLICANT: Chen, Rui-hong
 ; APPLICANT: Qian, Xiaohong B.
 ; APPLICANT: Wang, Zhiwei
 ; APPLICANT: Wehrman, Tom
 ; APPLICANT: Zhang, Jie
 ; APPLICANT: Zhou, Ping
 ; APPLICANT: Cao, Yi-Cheng
 ; APPLICANT: Drmanac, Radjoe T.
 ; TITLE OF INVENTION: Novel Nucleic Acids and Polypeptides
 ; FILE REFERENCE: 785CIP4CN
 ; CURRENT APPLICATION NUMBER: US/11/000,463
 ; CURRENT FILING DATE: 2004-11-29
 ; PRIOR FILING DATE: 2002-11-08
 ; PRIOR FILING DATE: 2001-01-25
 ; PRIOR FILING DATE: 2001-01-25
 ; PRIOR FILING DATE: 2001-01-25
 ; PRIOR FILING DATE: 2001-08-03
 ; PRIOR FILING DATE: 2000-01-25
 ; PRIOR FILING DATE: 2000-07-17
 ; PRIOR FILING DATE: 2000-07-17
 ; PRIOR FILING DATE: 2000-08-03
 ; PRIOR FILING DATE: 2000-08-03
 ; PRIOR FILING DATE: 2000-09-15
 ; NUMBER OF SEQ ID NOS: 944
 ; SOFTWARE: FastSeq for Windows Version 3.0
 ; SEQ ID NO 454
 ; LENGTH: 3674
 ; TYPE: PRT
 ; ORGANISM: Homo sapiens
 US-11-000-463-454

Query Match 46.8%; Score 44; DB 7; Length 3674;
 Best Local Similarity 45.0%; Pred. No. 57;
 Matches 9; Conservative 4; Mismatches 7; Indels 0; Gaps 0;

QY 1 QAQVNTIQGVSAALQTQQAE 20
 DB 1074 QQQVETLQGLLKQVQEQVAQ 1093

RESULT 15
 ; Sequence 58, Application US/11069934
 ; Publication No. US20050276811A1

```

; GENERAL INFORMATION:
; APPLICANT: CARROLL, MICHAEL C.
; APPLICANT: MOORE JR., FRANCIS D.
; APPLICANT: HECHTMAN, HERBERT B.
; TITLE OF INVENTION: NATURAL IGM ANTIBODIES AND INHIBITORS THEREOF
; FILE REFERENCE: CRA-002.01
; CURRENT APPLICATION NUMBER: US/11/069,834
; CURRENT FILING DATE: 2005-03-01
; PRIOR APPLICATION NUMBER: 60/588,648
; PRIOR FILING DATE: 2004-07-16
; PRIOR APPLICATION NUMBER: 60/549,123
; PRIOR FILING DATE: 2004-03-01
; NUMBER OF SEQ ID NOS: 65
; SOFTWARE: PatentIn Ver. 3.3
; SEQ ID NO 58
; LENGTH: 1992
; TYPE: PRT
; ORGANISM: Mus musculus
US-11-069-834-58

```

```

Query Match      44.7%; Score 42; DB 7; Length 1992;
Best Local Similarity 35.0%; Pred. No. 60;
Matches 7; Conservative 9; Mismatches 4; Indels 0; Gaps 0;

```

```

QY      1  QAOVNTIGQVSALQTQQA 20
Db      1243  EAFVSELKAEKLSLQTSR 1262

```

Search completed: January 25, 2006, 20:11:09
Job time : 6.5 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2006 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: January 25, 2006, 18:48:24 ; Search time 68.2 Seconds
(without alignments)
128.850 Million cell updates/sec

Title: US-10-797-821-10
Perfect score: 92
Sequence: 1 QQIQLSLSSKIVARNESLKQQ 20

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 2443163 seqs, 439378781 residues

Total number of hits satisfying chosen parameters: 2443163

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : A_Geneseq_21.*
1: Geneseqp1980s.*
2: Geneseqp1990s.*
3: Geneseqp2000s.*
4: Geneseqp2001s.*
5: Geneseqp2002s.*
6: Geneseqp2003as.*
7: Geneseqp2003Bs.*
8: Geneseqp2004s.*
9: Geneseqp2005s.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	92	100.0	20	7 ADD93630	Adx37253 Streptoco
2	92	100.0	20	7 ADD93630	Adx37253 Streptoco
3	92	100.0	431	7 ADD93649	Adx36649 Streptoco
4	92	100.0	431	7 ADD93650	Adx36650 Streptoco
5	92	100.0	431	7 ADD93653	Adx36653 Streptoco
6	92	100.0	431	9 ADX37272	Adx37272 Streptoco
7	92	100.0	431	9 ADX37273	Adx37273 Streptoco
8	92	100.0	431	9 ADX37276	Adx37276 Streptoco
9	92	100.0	431	9 AEB91500	Aeb91500 Microbial
10	92	100.0	432	7 ADD93651	Adx36651 Streptoco
11	92	100.0	432	7 ADD93652	Adx36652 Streptoco
12	92	100.0	432	9 ADX37274	Adx37274 Streptoco
13	92	100.0	432	9 ADX37275	Adx37275 Streptoco
14	82	89.1	395	5 ABP25918	Abp25918 Streptoco
15	82	89.1	447	5 ABP29684	Abp29684 Streptoco
16	82	89.1	447	8 ADU69524	Adu69524 S aglact
17	82	89.1	447	8 ADV88392	Adv88392 Streptoco
18	82	89.1	447	8 ADV81808	Adv81808 Streptoco
19	82	89.1	447	8 ADV79645	Adv79645 Streptoco
20	64	69.6	398	5 ABP25919	Abp25919 Streptoco
21	64	69.6	398	8 ADR83884	Adr83884 S. pyogen
22	64	69.6	398	9 AEB91542	Aeb91542 Microbial
23	49	53.3	210	2 Aay22579	Aay22579 Bacterial
24	49	53.3	392	6 ABU02747	Abu02747 S. pneumo

ALIGNMENTS

RESULT 1
ADD93630
ID ADD93630 standard; peptide; 20 AA.
AC ADD93630;
XX
XX
DT 29-JAN-2004 (first entry)
XX
DE Streptococcus mutans glucan binding protein-B peptide fragment.
XX
KW Glucan binding protein-B; GbpB; vaccine; anticaries; epitope; immunogen.
XX
OS Streptococcus mutans.
XX
PN WO2003075845-A2.
XX
PD 18-SEP-2003.
XX
PF 07-MAR-2003; 2003WO-US006962.
XX
PR 07-MAR-2002; 2002US-0363209P.
PR 08-AUG-2002; 2002US-0402483P.
XX
(FORS-) FORSYTH INST.
PA Smith DJ, Taubman MA;
PI
XX
DR WPI; 2003-845091/78.
XX
Composition useful as vaccines for dental caries comprises a fragment of a glucan binding protein-B binding to a major histocompatibility complex class II protein.
XX
PS Claim 11; Page 10; 49pp; English.
XX
The present sequence is that of a peptide comprising amino acid residues 88-107 of the glucan binding protein-B (GbpB) of Streptococcus mutans strain SK32 ADD93649. The peptide binds to a major histocompatibility complex (MHC) class II protein. It was identified as a potential B cell epitope using a matrix-based algorithm for epitope prediction, which was used to search the primary amino acid sequence of GbpB for known MHC class II binding motifs. The peptide can be used in immunogenic compositions and subunit vaccines for dental caries. These comprise MHC class II protein-binding GbpB peptides covalently linked with peptide subunits (preferably from the catalytic domain) of a glucosyltransferase. The compositions are used in a claimed method of eliciting production of an antibody in a mammal. Diepitopic or multiepitopic polypeptides can be

25 49 53.3 392 8 ADK47859 Adk47859 Streptoco
26 49 53.3 392 8 ADT50227 Adt50227 S. pneumon
27 49 53.3 392 8 ADT50226 Adt50226 S. pneumon
28 49 53.3 392 8 ADT50165 Adt50165 S. pneumon
29 49 53.3 399 8 ADR94595 ADR94595 Novel S.
30 49 53.3 399 9 AEA58465 Aea58465 Streptoco
31 49 53.3 415 8 ADG76637 Adg76637 Phosphoil
32 49 53.3 578 4 ABB71138 Abb71138 Drosophil
33 47.5 51.6 281 8 ADS11023 Ads11023 Human the
34 47 51.1 422 2 AAY00251 Aay00251 Enterococ
35 47 51.1 422 5 ABP43470 Abp43470 E faecali
36 47 51.1 422 6 ABU88498 Abu88498 E. faecal
37 47 51.1 422 6 ABUL3749 Abul3749 Enterococ
38 47 51.1 422 9 ADY39218 Ady39218 Novel Ent
39 47 51.1 449 2 AAY00250 Aay00250 Enterococ
40 47 51.1 449 5 ABP43469 Abp43469 E faecali
41 47 51.1 449 6 ABU88497 Abu88497 E. faecal
42 47 51.1 449 6 ABUL3748 Abul3748 Enterococ
43 47 51.1 449 9 ADV16734 Adv16734 E. faecal
44 47 51.1 449 9 ADY39216 Ady39216 Novel Ent
45 47 51.1 450 7 ADH87829 Adh87829 Enterococ

CC prepared synthetically or by recombinant DNA technology. Antibodies
CC raised against MHC class II binding fragments of GbpB can be used in
CC passive immunisation.

SQ Sequence 20 AA;

Query Match 100.0%; Score 92; DB 7; Length 20;
Best Local Similarity 100.0%; Pred. No. 1.3e-08;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 QOIQTLSKIVARNESLKQ 20

Db 1 QOIQTLSKIVARNESLKQ 20

RESULT 2

ADX37253
ID ADX37253 standard; peptide; 20 AA.

AC ADX37253;

DT 21-APR-2005 (first entry)

DE Streptococcus mutant glucan binding protein B peptide #10.

KW immunogenicity; immune stimulation; glucan binding protein-B;
KW microparticle; major histocompatibility complex; tooth disease.

OS Streptococcus mutans.

PN US2005031633-A1.

PD 10-FEB-2005.

PF 09-MAR-2004; 2004US-00797821.

PR 13-APR-1998; 98US-0081550P.

PR 08-JAN-1999; 99US-0115142P.

PR 12-APR-1999; 99US-00290049.

PR 07-MAR-2002; 2002US-0363209P.

PR 08-AUG-2002; 2002US-0402483P.

PR 07-MAR-2003; 2003US-00383930.

PA (SMIT/) SMITH D J.

PI (TAUB/) TAUBMAN M A.

PI Smith DJ, Taubman MA;

PI WPI; 2005-151644/16.

PS Claim 4; SEQ ID NO 10; 73pp; English.

XX The invention relates to a composition comprising a fragment of a glucan
XX binding protein-B (GbpB) and a biocompatible microparticle, where the
XX fragment binds to a major histocompatibility complex (MHC) class II
XX protein. The composition is useful for producing an antibody for
XX immunizing mammals against dental caries. This sequence corresponds to a
XX Streptococcus mutans GbpB-derived peptide of the invention.

SQ Sequence 20 AA;

Query Match 100.0%; Score 92; DB 9; Length 20;
Best Local Similarity 100.0%; Pred. No. 1.3e-08;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 QOIQTLSKIVARNESLKQ 20

Db 1 QOIQTLSKIVARNESLKQ 20

RESULT 3
ADD93649
ID ADD93649 standard; protein; 431 AA.
XX
AC ADD93649;
XX
DT 29-JAN-2004 (first entry)
XX
DE Streptococcus mutans glucan binding protein-B.
XX
KW Glucan binding protein-B; GbpB; vaccine; anticonaries; epitope; immunogen.
XX
OS Streptococcus mutans.
XX
FH Key Location/Qualifiers
FT Region 6..25
FT /note= "HLA-binding peptide"
FT Region 16..35
FT /note= "HLA-binding peptide"
FT Region 33..52
FT /note= "HLA-binding peptide"
FT Region 37..56
FT /note= "HLA-binding peptide"
FT Region 48..67
FT /note= "HLA-binding peptide"
FT Region 52..71
FT /note= "HLA-binding peptide"
FT Region 88..107
FT /note= "HLA-binding peptide"
FT Region 113..132
FT /note= "HLA-binding peptide"
FT Region 117..136
FT /note= "HLA-binding peptide"
FT Region 137..156
FT /note= "HLA-binding peptide"
FT Region 174..193
FT /note= "HLA-binding peptide"
FT Region 194..213
FT /note= "HLA-binding peptide"
FT Region 214..233
FT /note= "HLA-binding peptide"
FT Region 248..267
FT /note= "HLA-binding peptide"
FT Region 289..308
FT /note= "HLA-binding peptide"
FT Region 306..325
FT /note= "HLA-binding peptide"
FT Region 311..330
FT /note= "HLA-binding peptide"
FT Region 349..368
FT /note= "HLA-binding peptide"
FT Region 365..384
FT /note= "HLA-binding peptide"
FT Region 383..402
FT /note= "HLA-binding peptide"
FT Region 403..422
FT /note= "HLA-binding peptide"
XX
PN WO2003075845-A2.
XX
PD 18-SEP-2003.
XX
PF 07-MAR-2003; 2003WO-US006962.
XX
PR 07-MAR-2002; 2002US-0363209P.
PR 08-AUG-2002; 2002US-0402483P.
XX
PA (FORS-) FORSYTH INST.
XX
PI Smith DJ, Taubman MA;
XX

Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 QQIQTLSKIVARNESLKQ 20
 Db 88 QQIQTLSKIVARNESLKQ 107

RESULT 6
 ADX37272
 ID ADX37272 standard; protein; 431 AA.
 XX AC ADX37272;
 XX DT 21-APR-2005 (first entry)
 XX DE Streptococcus mutant glucan binding protein B variant #1.
 XX KW immunogenicity; immune stimulation; glucan binding protein-B;
 KW microparticle; major histocompatibility complex; tooth disease.
 XX OS Streptococcus mutans.
 XX PN US2005031633-A1.
 XX PD 10-FEB-2005.
 XX PF 09-MAR-2004; 2004US-00797821.
 XX PR 13-APR-1998; 98US-0081550P.
 PR 08-JAN-1999; 99US-0115142P.
 PR 12-APR-1999; 99US-00290049.
 PR 07-MAR-2002; 2002US-0363209P.
 PR 08-AUG-2002; 2002US-0402483P.
 PR 07-MAR-2003; 2003US-00383930.
 XX (SMIT/) SMITH D J.
 PA (TAUB/) TAUBMAN M A.
 XX Smith DJ, Taubman MA;
 PI WPI; 2005-151644/16.
 XX New composition comprising a fragment of a glucan binding protein-B
 (GbpB) that binds to MHC class II protein, and a biocompatible
 microparticle, useful for producing an antibody (claimed) for immunizing
 mammals against dental caries.
 XX Claim 3; SEQ ID NO 29; 73pp; English.

QY The invention relates to a composition comprising a fragment of a glucan
 binding protein-B (GbpB) and a biocompatible microparticle, where the
 fragment binds to a major histocompatibility complex (MHC) class II
 protein. The composition is useful for producing an antibody for
 immunizing mammals against dental caries. This sequence corresponds to a
 Streptococcus mutans GbpB protein of the invention.

XX Sequence 431 AA;
 Query Match 100.0%; Score 92; DB 9; Length 431;
 Best Local Similarity 100.0%; Pred. No. 4.3e-07;
 Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 QQIQTLSKIVARNESLKQ 20
 Db 88 QQIQTLSKIVARNESLKQ 107

RESULT 7
 ADX37273
 ID ADX37273 standard; protein; 431 AA.
 XX AC ADX37273;
 XX DT 21-APR-2005 (first entry)
 XX DE Streptococcus mutant glucan binding protein B variant #5.
 XX KW immunogenicity; immune stimulation; glucan binding protein-B;
 KW microparticle; major histocompatibility complex; tooth disease.
 XX OS Streptococcus mutans.
 XX PN US2005031633-A1.
 XX PD 10-FEB-2005.

DT 21-APR-2005 (first entry)
 XX Streptococcus mutant glucan binding protein B variant #2.
 DE immunogenicity; immune stimulation; glucan binding protein-B;
 KW microparticle; major histocompatibility complex; tooth disease.
 XX OS Streptococcus mutans.
 XX PN US2005031633-A1.
 XX PD 10-FEB-2005.
 XX PF 09-MAR-2004; 2004US-00797821.
 XX PR 13-APR-1998; 98US-0081550P.
 PR 08-JAN-1999; 99US-0115142P.
 PR 12-APR-1999; 99US-00290049.
 PR 07-MAR-2002; 2002US-0363209P.
 PR 08-AUG-2002; 2002US-0402483P.
 PR 07-MAR-2003; 2003US-00383930.
 XX (SMIT/) SMITH D J.
 PA (TAUB/) TAUBMAN M A.
 XX Smith DJ, Taubman MA;
 PI WPI; 2005-151644/16.
 XX New composition comprising a fragment of a glucan binding protein-B
 (GbpB) that binds to MHC class II protein, and a biocompatible
 microparticle, useful for producing an antibody (claimed) for immunizing
 mammals against dental caries.
 XX Claim 3; SEQ ID NO 30; 73pp; English.

XX The invention relates to a composition comprising a fragment of a glucan
 binding protein-B (GbpB) and a biocompatible microparticle, where the
 fragment binds to a major histocompatibility complex (MHC) class II
 protein. The composition is useful for producing an antibody for
 immunizing mammals against dental caries. This sequence corresponds to a
 Streptococcus mutans GbpB protein of the invention.

XX Sequence 431 AA;
 Query Match 100.0%; Score 92; DB 9; Length 431;
 Best Local Similarity 100.0%; Pred. No. 4.3e-07;
 Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 QQIQTLSKIVARNESLKQ 20
 Db 88 QQIQTLSKIVARNESLKQ 107

RESULT 8
 ADX37276
 ID ADX37276 standard; protein; 431 AA.
 XX AC ADX37276;
 XX DT 21-APR-2005 (first entry)
 XX DE Streptococcus mutant glucan binding protein B variant #5.
 XX KW immunogenicity; immune stimulation; glucan binding protein-B;
 KW microparticle; major histocompatibility complex; tooth disease.
 XX OS Streptococcus mutans.
 XX PN US2005031633-A1.
 XX PD 10-FEB-2005.

PF 09-MAR-2004; 2004US-00797821.
 XX 13-APR-1998; 98US-0081550P.
 PR 08-JAN-1999; 99US-0115142P.
 PR 12-APR-1999; 99US-00290049.
 PR 07-MAR-2002; 2002US-0363209P.
 PR 08-AUG-2002; 2002US-0402483P.
 PR 07-MAR-2003; 2003US-00383930.
 XX (SMIT/) SMITH D J.
 PA (TAUB/) TAUBMAN M A.
 XX Smith DJ, Taubman MA;
 PI WPI; 2005-151644/16.
 DR
 XX New composition comprising a fragment of a glucan binding protein-B
 PT (GbpB) that binds to MHC class II protein, and a biocompatible
 PT microparticle, useful for producing an antibody (claimed) for immunizing
 PT mammals against dental caries.
 XX Claim 3; SEQ ID NO 33; 73pp; English.
 PS The invention relates to a composition comprising a fragment of a glucan
 CC binding protein-B (GbpB) and a biocompatible microparticle, where the
 CC fragment binds to a major histocompatibility complex (MHC) class II
 CC protein. The composition is useful for producing an antibody for
 CC immunizing mammals against dental caries. This sequence corresponds to a
 CC Streptococcus mutans GbpB protein of the invention.
 XX Sequence 431 AA;
 SQ
 Query Match 100.0%; Score 92; DB 9; Length 431;
 Best Local Similarity 100.0%; Pred. No. 4.3e-07;
 Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 QQIQTLLSSKIVARNESLKKQ 20
 Db 88 QQIQTLLSSKIVARNESLKKQ 107
 RESULT 9
 AEB91500
 ID AEB91500 standard; protein; 431 AA.
 AC AEB91500;
 XX 20-OCT-2005 (first entry)
 DT
 XX Microbial pathogen adhesin protein sequence, SEQ ID NO:210.
 DE algorithm; adhesin; pharmaceutical; vaccine; drug screening;
 KW bordetella pertussis infection; antibacterial; pneumonia;
 KW antiinflammatory; respiratory-gen.; gastric ulcer; antiulcer;
 KW gastrointestinal-gen.; urinary tract infection; antimicrobial; uropathic.
 XX Streptococcus mutans.
 OS WO2005076010-A2.
 XX 18-AUG-2005.
 PD
 XX 07-FEB-2005; 2005WO-IN000037.
 PF
 XX 06-FEB-2004; 2004IN-DE000173.
 PR 20-JUL-2004; 2004US-0589227P.
 PR
 XX (COUL) COUNCIL SCI & IND RES SOUTH AFRICA.
 PA Sachdeva G, Kumar K, Jain P, Brahmachari SK, Ramachandran S;
 XX WPI; 2005-597835/61.
 DR
 XX

PT Computational method for identifying adhesin and adhesin like molecules,
 PT comprises computing sequence-based attributes of protein sequences using
 PT neural network software and training an artificial neural network.
 XX Claim 16; SEQ ID NO 210; 402pp; English.
 XX The present invention relates to a computational method (M1) for
 CC identifying adhesin and adhesin-like proteins, by computing the sequence-
 CC based attributes of protein sequences using five attribute modules of a
 CC neural network software, training an artificial neural network (ANN) for
 CC each of the computed five attributes, and identifying the adhesin and
 CC adhesin-like proteins having probability of being an adhesin (Pad) as
 CC equal or greater than 0.51. Also claimed is a set of 274 annotated genes
 CC encoding adhesin and adhesin-like proteins, having 274 fully defined 162-
 CC 11151 base pair (SEQ ID NO: 385-658) sequences; a set of 105 hypothetical
 CC genes encoding adhesin and adhesin-like proteins, having 105 fully
 CC defined 306-15876 base pairs (SEQ ID NO: 659-763); a set of 279 annotated
 CC adhesin and adhesin-like proteins, having 279 fully defined 53-3716 base
 CC pair (SEQ ID NO: 1-279); a set of 105 hypothetical adhesin and adhesin-
 CC like proteins, having 105 fully defined 106-5291 base pair (SEQ ID NO:
 CC 280-384) sequences; and a fully connected multilayer feed forward ANN (I)
 CC based on (M1). (M1) is useful for identifying adhesin and adhesin-like
 CC proteins, of therapeutic potential, and identifying and short-listing
 CC to eliminate diseases caused by various pathogenic organisms. (M1) is
 CC useful for identifying putative adhesins that are important in drug
 CC discovery and preventing therapeutics for whooping cough, pneumonia,
 CC gastric ulcer and urinary tract infections. (M1) identifies adhesins from
 CC distantly related organisms, and from bacteria belonging to a wide
 CC phylogenetic spectrum. (M1) is capable of predicting adhesive nature of
 CC unique proteins. The present sequence is a microbial pathogen adhesin
 CC protein sequence.
 XX Sequence 431 AA;
 SQ
 Query Match 100.0%; Score 92; DB 9; Length 431;
 Best Local Similarity 100.0%; Pred. No. 4.3e-07;
 Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 QQIQTLLSSKIVARNESLKKQ 20
 Db 88 QQIQTLLSSKIVARNESLKKQ 107
 RESULT 10
 ADD93651
 ID ADD93651 standard; protein; 432 AA.
 XX ADD93651;
 AC ADD93651;
 XX 29-JAN-2004 (first entry)
 DT Streptococcus mutans glucan binding protein-B.
 DE Glucan binding protein-B; GbpB; vaccine; anticaries; epitope; immunogen.
 XX Streptococcus mutans.
 OS WO2003075845-A2.
 XX 18-SEP-2003.
 PD
 XX 07-MAR-2003; 2003WO-US006962.
 PF
 XX 07-MAR-2002; 2002US-0363209P.
 PR 08-AUG-2002; 2002US-0402483P.
 PR
 XX (FORS-) FORSYTH INST.
 PA Smith DJ, Taubman MA;
 XX WPI; 2003-845091/78.
 DR GENBANK; AY046412.
 DR

XX Composition useful as vaccines for dental caries comprises a fragment of
PT a glucan binding protein-B binding to a major histocompatibility complex
PT class II protein.
XX
XX Claim 5; Page 8; 49pp; English.
XX
XX The present sequence is the protein sequence of the glucan binding
CC protein-B (GbpB) of Streptococcus mutans strain 15JF2. The sequence
CC includes a number of human leucocyte antigen (HLA)-binding epitopes. The
CC invention provides immunogenic compositions and vaccines for dental
CC caries. The compositions comprise major histocompatibility complex (MHC)
CC class II protein-binding GbpB peptides, especially HLA-binding peptides,
CC covalently linked with peptide subunits of a glucosyltransferase. The
CC compositions are used in a claimed method of eliciting production of an
CC antibody in a mammal. Diepitopic or multiepitopic polypeptides can be
CC prepared synthetically or by recombinant DNA technology. Antibodies
CC raised against MHC class II binding fragments of GbpB can be used in
CC passive immunisation.
XX
XX SQ Sequence 432 AA;

XX Query Match 100.0%; Score 92; DB 7; Length 432;
XX Best Local Similarity 100.0%; Pred. No. 4.3e-07;
XX Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

XX QY 1 QQIQTLSSKIVARNESLKQ 20
XX |||||
XX DB 88 QQIQTLSSKIVARNESLKQ 107

XX RESULT 11
XX ADD93652
XX ID ADD93652 standard; protein; 432 AA.
XX AC ADD93652;
XX
XX 29-JAN-2004 (first entry)
XX
XX Streptococcus mutans glucan binding protein-B.
XX
XX Glucan binding protein-B; GbpB; vaccine; anticaries; epitope; immunogen.
XX
XX Streptococcus mutans.
XX
XX WO2003075845-A2.
XX
XX 18-SEP-2003.
XX
XX 07-MAR-2003; 2003WO-US006962.
XX
XX 07-MAR-2002; 2002US-0363209P.
XX
XX 08-AUG-2002; 2002US-0402483P.
XX
XX (FORS-) FORSYTH INST.
XX
XX Smith DJ, Taubman MA;
XX
XX WPI; 2003-845091/78.
XX GENBANK; AY046413.
XX
XX Composition useful as vaccines for dental caries comprises a fragment of
PT a glucan binding protein-B binding to a major histocompatibility complex
PT class II protein.
XX
XX Claim 5; Page 8; 49pp; English.
XX
XX The present sequence is the protein sequence of the glucan binding
CC protein-B (GbpB) of Streptococcus mutans strain 3S1. The sequence
CC includes a number of human leucocyte antigen (HLA)-binding epitopes. The
CC invention provides immunogenic compositions and vaccines for dental
CC caries. The compositions comprise major histocompatibility complex (MHC)
CC class II protein-binding GbpB peptides, especially HLA-binding peptides,
CC covalently linked with peptide subunits of a glucosyltransferase. The
CC compositions are used in a claimed method of eliciting production of an
CC antibody in a mammal. Diepitopic or multiepitopic polypeptides can be
CC prepared synthetically or by recombinant DNA technology. Antibodies
CC raised against MHC class II binding fragments of GbpB can be used in
CC passive immunisation.
XX

CC covalently linked with peptide subunits of a glucosyltransferase. The
CC compositions are used in a claimed method of eliciting production of an
CC antibody in a mammal. Diepitopic or multiepitopic polypeptides can be
CC prepared synthetically or by recombinant DNA technology. Antibodies
CC raised against MHC class II binding fragments of GbpB can be used in
CC passive immunisation.
XX
XX SQ Sequence 432 AA;

XX Query Match 100.0%; Score 92; DB 7; Length 432;
XX Best Local Similarity 100.0%; Pred. No. 4.3e-07;
XX Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

XX QY 1 QQIQTLSSKIVARNESLKQ 20
XX |||||
XX DB 88 QQIQTLSSKIVARNESLKQ 107

XX RESULT 12
XX ADX37274
XX ID ADX37274 standard; protein; 432 AA.
XX AC ADX37274;
XX
XX 21-APR-2005 (first entry)
XX
XX Streptococcus mutant glucan binding protein B variant #3.
XX
XX immunogenicity; immune stimulation; glucan binding protein-B;
XX microparticle; major histocompatibility complex; tooth disease.
XX
XX Streptococcus mutans.
XX
XX US2005031633-A1.
XX
XX 10-FEB-2005.
XX
XX 09-MAR-2004; 2004US-00797821.
XX
XX 13-APR-1998; 98US-0081550P.
XX
XX 08-JAN-1999; 99US-0115142P.
XX
XX 12-APR-1999; 99US-0029004P.
XX
XX 07-MAR-2002; 2002US-0363209P.
XX
XX 08-AUG-2002; 2002US-0402483P.
XX
XX 07-MAR-2003; 2003US-00383930.
XX
XX (SMIT/) SMITH D J.
XX (TAUB/) TAUBMAN M A.
XX
XX Smith DJ, Taubman MA;
XX
XX WPI; 2005-151644/16.
XX
XX New composition comprising a fragment of a glucan binding protein-B
XX (GbpB) that binds to MHC class II protein, and a biocompatible
XX microparticle, useful for producing an antibody (claimed) for immunizing
XX mammals against dental caries.
XX
XX Claim 3; SEQ ID NO 31; 73pp; English.
XX
XX The invention relates to a composition comprising a fragment of a glucan
CC binding protein-B (GbpB) and a biocompatible microparticle, where the
CC fragment binds to a major histocompatibility complex (MHC) class II
CC protein. The composition is useful for producing an antibody for
CC immunizing mammals against dental caries. This sequence corresponds to a
CC Streptococcus mutans GbpB protein of the invention.
XX
XX SQ Sequence 432 AA;

XX Query Match 100.0%; Score 92; DB 9; Length 432;
XX Best Local Similarity 100.0%; Pred. No. 4.3e-07;
XX Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 QOIQTLSKIVARNESLKQ 20
 Db 88 QOIQTLSKIVARNESLKQ 107

RESULT 13
 ADX37275
 ID ADX37275 standard; protein; 432 AA.
 AC ADX37275;
 XX
 DT 21-APR-2005 (first entry)
 XX
 DE Streptococcus mutant glucan binding protein B variant #4.
 XX
 KW immunogenicity; immune stimulation; glucan binding protein-B;
 KW microparticle; major histocompatibility complex; tooth disease.
 XX
 OS Streptococcus mutans.
 XX
 PN US2005031633-A1.
 XX
 PD 10-FEB-2005.
 XX
 PF 09-MAR-2004; 2004US-00797821.
 XX
 PR 13-APR-1998; 98US-0081550P.
 PR 08-JAN-1999; 99US-0115142P.
 PR 12-APR-1999; 99US-00290049.
 PR 07-MAR-2002; 2002US-0363209P.
 PR 08-AUG-2002; 2002US-0402483P.
 PR 07-MAR-2003; 2003US-00383930.
 XX
 PA (SMIT/) SMITH D J.
 PA (TAUB/) TAUBMAN M A.
 XX
 PI Smith DJ, Taubman MA;
 XX
 DR WPI; 2005-151644/16.
 XX
 PT New composition comprising a fragment of a glucan binding protein-B
 PT (GbpB) that binds to MHC class II protein, and a biocompatible
 PT microparticle, useful for producing an antibody (claimed) for immunizing
 PT mammals against dental caries.
 XX
 PS Claim 3; SEQ ID NO 32; 73pp; English.
 XX
 CC The invention relates to a composition comprising a fragment of a glucan
 CC binding protein-B (GbpB) and a biocompatible microparticle, where the
 CC fragment binds to a major histocompatibility complex (MHC) class II
 CC protein. The composition is useful for producing an antibody for
 CC immunizing mammals against dental caries. This sequence corresponds to a
 CC Streptococcus mutans GbpB protein of the invention.
 XX
 SQ Sequence 432 AA;

Query Match 100.0%; Score 92; DB 9; Length 432;
 Best Local Similarity 100.0%; Pred. No. 4.3e-07;
 Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 QOIQTLSKIVARNESLKQ 20
 Db 88 QOIQTLSKIVARNESLKQ 107

RESULT 14
 ABP25918
 ID ABP25918 standard; protein; 395 AA.
 XX
 AC ABP25918;
 XX
 DT 02-JUL-2002 (first entry)
 XX
 DE Streptococcus polypeptide SEQ ID NO 8544.
 XX
 KW Streptococcus; GAS; GBS; group B streptococcus; Streptococcus agalactiae;

DE Streptococcus polypeptide SEQ ID NO 1012.
 XX Streptococcus; GAS; GBS; group B streptococcus; Streptococcus agalactiae;
 KW group A streptococcus; Streptococcus pyogenes; antibacterial;
 KW antiinflammatory; infection; vaccine; meningitis; gene therapy.
 XX Streptococcus agalactiae.
 OS Streptococcus agalactiae.
 XX
 PN WO200234771-A2.
 XX
 PD 02-MAY-2002.
 XX
 PF 29-OCT-2001; 2001WO-GB004789.
 XX
 PR 27-OCT-2000; 2000GB-00026333.
 PR 24-NOV-2000; 2000GB-00028727.
 PR 07-MAR-2001; 2001GB-00005640.
 XX
 PA (CHIR-) CHIRON SPA.
 PA (GENO-) INST GENOMIC RES.
 XX
 PI Telford J, Maignani V, Margarit Y Rosi, Grandi G, Fraser C;
 PI Tettelin H;
 XX
 DR WPI; 2002-352536/38.
 DR N-PSDB; ABN66549.
 XX
 PT New Streptococcus protein for the treatment or prevention of infection or
 PT disease caused by Streptococcus bacteria, such as meningitis, and for
 PT detecting a compound that binds to the protein.
 XX
 PS Claim 1; Page 3256; 4525pp; English.
 XX
 CC The invention relates to a protein (ABP25413-ABP30895) from group B
 CC streptococcus/GBS (Streptococcus agalactiae) or group A streptococcus/GAS
 CC (Streptococcus pyogenes), comprising one of 5483 sequences (SI), given in
 CC the specification. The proteins have antibacterial and antiinflammatory
 CC activity. (I), nucleic acids encoding (I), ABN66044-ABN71526 and
 CC antibodies that bind (I) are used in the manufacture of medicaments for
 CC the treatment or prevention of infection or disease caused by
 CC Streptococcus bacteria, particularly S. agalactiae and S. pyogenes.
 CC Nucleic acids encoding (I) are used to detect Streptococcus in a
 CC biological sample. (I) is used to determine whether a compound binds to
 CC (I). A composition comprising (I) or a nucleic acid encoding (I), may be
 CC used as a vaccine or diagnostic composition. The disease caused by
 CC Streptococcus that is prevented or treated may be meningitis. Nucleic
 CC acid encoding (I) may be used to recombinantly produce (I) and may be
 CC used in gene therapy. Antibodies to (I) are used for affinity
 CC chromatography, immunoassays, and distinguishing/identifying
 CC Streptococcus proteins
 XX
 SQ Sequence 395 AA;

Query Match 89.1%; Score 82; DB 5; Length 395;
 Best Local Similarity 85.0%; Pred. No. 2.1e-05;
 Matches 17; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

QY 1 QOIQTLSKIVARNESLKQ 20
 Db 34 QEIQTLSKIVARNESLKQ 53

RESULT 15
 ABP29684
 ID ABP29684 standard; protein; 447 AA.
 XX
 AC ABP29684;
 XX
 DT 02-JUL-2002 (first entry)
 XX
 DE Streptococcus polypeptide SEQ ID NO 8544.
 XX
 KW Streptococcus; GAS; GBS; group B streptococcus; Streptococcus agalactiae;

KW group A streptococcus; Streptococcus pyogenes; antibacterial;
 KW antiinflammatory; infection; vaccine; meningitis; gene therapy.

OS Streptococcus agalactiae.

PN WO200234771-A2.

XX

PD 02-MAY-2002.

XX 29-OCT-2001; 2001WO-GB004789.

XX 27-OCT-2000; 2000GB-00026333.

PR 24-NOV-2000; 2000GB-00028727.

PR 07-MAR-2001; 2001GB-00005640.

XX

XX (CHIR-) CHIRON SPA.

PA (GENO-) INST GENOMIC RES.

XX

XX Telford J, Maignani V, Margarit Y Rosi, Grandi G, Fraser C;

PI Tettelin H;

XX

DR WPI; 2002-352536/38.

DR N-PSDB; ABN70315.

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

STREPTOCOCCUS AGALACTIAE (GROUP B) IS A CAUSE OF INFECTION OR DISEASE CAUSED BY STREPTOCOCCUS BACTERIA, SUCH AS MENINGITIS, AND FOR DETECTING A COMPOUND THAT BINDS TO THE PROTEIN.

Claim 1; Page 3965; 4525pp; English.

The invention relates to a protein (ABP25413-ABP30895) from group B streptococcus/GBS (Streptococcus agalactiae) or group A streptococcus/GAS (Streptococcus pyogenes), comprising one of 5483 sequences (S1), given in the specification. The proteins have antibacterial and antiinflammatory activity. (I), nucleic acids encoding (I), ABN66044-ABN71526 and antibodies that bind (I) are used in the manufacture of medicaments for the treatment or prevention of infection or disease caused by Streptococcus bacteria, particularly S. agalactiae and S. pyogenes. Nucleic acids encoding (I) are used to detect Streptococcus in a biological sample. (I) is used to determine whether a compound binds to (I). A composition comprising (I) or a nucleic acid encoding (I), may be used as a vaccine or diagnostic composition. The disease caused by Streptococcus that is prevented or treated may be meningitis. Nucleic acid encoding (I) may be used to recombinantly produce (I) and may be used in gene therapy. Antibodies to (I) are used for affinity chromatography, immunoassays, and distinguishing/identifying Streptococcus proteins

Sequence 447 AA;

Query Match 89.1%; Score 82; DB 5; Length 447;
 Best Local Similarity 85.0%; Pred. No. 2.4e-05;
 Matches 17; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

QY 1 QQIQTLSSKIVARNESLKKQ 20
 Db 86 QEIQTLSNKNIVARNESLKKQ 105

Search completed: January 25, 2006, 19:01:45
 Job time : 68.2 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2006 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: January 25, 2006, 18:48:55 ; Search time 9.1 Seconds
(without alignments)
211.465 Million cell updates/sec

Title: US-10-797-821-10
Perfect score: 92
Sequence: 1 QIQTLSSKIVARNESLKQ 20

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283416 seqs, 96216763 residues
Total number of hits satisfying chosen parameters: 283416

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : PIR 80:*
1: pir1:*
2: pir2:*
3: pir3:*
4: pir4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	49	53.3	392	2 G95258	secreted 45 kd pro
2	49	53.3	392	2 B98124	general stress pro
3	46	50.0	1185	2 I58009	gene ERCC5 protein
4	46	50.0	1186	2 S35993	DNA repair protein
5	46	50.0	1690	2 T13030	microtubule bindin
6	45	48.9	169	2 D86801	prophage p13 prote
7	45	48.9	418	2 H82485	thermolabile hemol
8	44	47.8	71	2 D82743	hypothetical prote
9	44	47.8	297	2 H72670	hypothetical prote
10	44	47.8	526	2 G83436	hypothetical prote
11	44	47.8	913	2 T15278	hypothetical prote
12	44	47.8	1780	2 T17272	hypothetical prote
13	44	47.8	4077	2 T17484	hypothetical prote
14	42	45.7	233	2 E86340	protein F2D10.32 l
15	42	45.7	329	2 T51853	RING-H2 finger pro
16	42	45.7	354	2 G71403	hypothetical prote
17	42	45.7	376	2 C85155	RING-H2 finger pro
18	42	45.7	458	2 S74941	hypothetical prote
19	42	45.7	741	2 QJ0375	hypothetical prote
20	42	45.7	979	2 E72236	clostrypain-relate
21	41	44.6	140	2 E82229	hypothetical prote
22	41	44.6	205	2 S60005	Mad3 protein - mou
23	41	44.6	211	2 E83089	ATP-phosphoribosyl
24	41	44.6	230	2 T51265	probable HLH DNA b
25	41	44.6	396	2 C97512	ggdef family prote
26	41	44.6	405	2 T19857	hypothetical prote
27	41	44.6	407	2 T14909	bZIP DNA-binding p
28	41	44.6	415	2 AD2731	GGDEF family prote
29	41	44.6	434	2 S42676	KES1 protein - yea

30	41	44.6	585	2 F90460	AAA family ATPase
31	41	44.6	803	2 T40514	Chaperonin hsp/8p
32	41	44.6	981	2 T05505	hypothetical prote
33	41	44.6	1786	1 MMHUB1	laminin beta-1 cha
34	41	44.6	2077	1 WZBEZ4	240K tegument prot
35	40.5	44.0	117	2 T46322	hypothetical prote
36	40.5	44.0	158	2 T40378	hypothetical prote
37	40.5	44.0	583	1 A41129	radixin - mouse
38	40.5	44.0	583	1 A46127	radixin - human
39	40.5	44.0	583	1 S39805	radixin - pig
40	40	43.5	110	2 S74519	suppressor protein
41	40	43.5	134	2 B83393	probable transcrip
42	40	43.5	174	2 C48583	stress-inducible p
43	40	43.5	204	2 S56256	hypothetical prote
44	40	43.5	222	2 G97129	hypothetical prote
45	40	43.5	253	2 T02106	hypothetical prote

ALIGNMENTS

RESULT 1

G95258
secreted 45 kd protein [imported] - Streptococcus pneumoniae (strain TIGR4)
C:Species: Streptococcus pneumoniae
C:Date: 03-Aug-2001 #sequence_revision 03-Aug-2001 #text_change 09-Jul-2004
C:Accession: G95258
R:Tettelin, H.; Neilson, K.E.; Paulsen, I.T.; Eisen, J.A.; Read, T.D.; Peterson, S.; Helon, J.D.; Umayam, L.A.; White, O.; Salzberg, S.L.; Lewis, M.R.; Radune, D.; Holtzapfel, T.; Hickey, E.K.; Holt, I.E.
Science 293, 498-506, 2001
A:Authors: Loftus, B.J.; Yang, F.; Smith, H.O.; Venter, J.C.; Dougherty, B.A.; Morrison A:Title: Complete Genome Sequence of a Virulent Isolate of Streptococcus pneumoniae.
A:Reference number: A95000; MUID:21357209; PMID:11463916
A:Accession: G95258
A>Status: Preliminary
A:Molecule type: DNA
A:Residues: 1-392 <KUR>
A:Cross-references: UNIPROT:Q97N55; UNIPROT:Q8DMY4; UNIPARC:UPI0000051B81; GB:AE005672;
A:Experimental source: strain TIGR4
C:Genetics:
A:Gene: SP2216

Query Match 53.3%; Score 49; DB 2; Length 392;
Best Local Similarity 52.6%; Pred. No. 3.3;
Matches 10; Conservative 5; Mismatches 4; Indels 0; Gaps 0;

QY 2 QIQTLSSKIVARNESLKQ 20
DB 89 EITELSKNIVSRNQSLKQ 107

RESULT 2

B98124
general stress protein GSP-781 [imported] - Streptococcus pneumoniae (strain R6)
C:Species: Streptococcus pneumoniae
C:Date: 22-Oct-2001 #sequence_revision 22-Oct-2001 #text_change 09-Jul-2004
C:Accession: B98124
R:Hoskins, J.A.; Alborn Jr., W.; Arnold, J.; Blaszcak, L.; Burgett, S.; DeHoff, B.S.; e, R.; LeBlanc, D.; Lee, L.N.; Lefkowitz, E.J.; Lu, J.; Matsushima, P.; McAhren, S.; y, P.; Sun, P.M.; Winkler, M.E.
J. Bacteriol. 183, 5709-5717, 2001
A:Authors: Yang, Y.; Young-Bellido, M.; Zhao, G.; Zook, C.; Baltz, R.H.; Jaskunas, S.R. A:Title: Genome of the Bacterium Streptococcus pneumoniae Strain R6.
A:Reference number: A97872; MUID:21429245; PMID:11544234
A:Accession: B98124
A>Status: Preliminary
A:Molecule type: DNA
A:Residues: 1-392 <KUR>
A:Cross-references: UNIPROT:Q97N55; UNIPROT:Q8DMY4; UNIPARC:UPI0000051B81; GB:AE007317;
C:Genetics:
A:Gene: gsp-781

Best Local Similarity 47.4%; Score 44; DB 2; Length 418; Mismatches 3; Conservative 9; Indels 7; Gaps 0;

Query 1 QOIQTLSKIVARNESLKQ 19
 Db 40 QNISILKQRLDRNEQLKQ 58

RESULT 7
 H82485
 thermolabile hemolysin VCA0218 [imported] - Vibrio cholerae (strain N16961 serogroup O1)
 C:Species: Vibrio cholerae
 C:Date: 18-Aug-2000 #sequence_revision 20-Aug-2000 #text_change 09-Jul-2004
 C:Accession: H82485
 R:Heidelberg, J.F.; Eisen, J.A.; Nelson, W.C.; Clayton, R.A.; Gwinn, M.L.; Dodson, R.J.; Chardson, D.; Smolaeava, M.D.; Vamathevan, J.; Baas, S.; Qin, H.; Dragoi, I.; Sellers, E. I., R.R.; Mekalanos, J.J.; Venter, J.C.; Fraser, C.M.
 Nature 406, 477-483, 2000
 A:Title: DNA Sequence of both chromosomes of the cholera pathogen Vibrio cholerae.
 A:Reference number: A82035; MUID:20406833; PMID:10952301
 A:Accession: H82485
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-418 <HEI>
 A:Cross-references: UNIPROT:Q9KMW0; UNIPARC:UPI00000C3482; GB:AE004362; GB:AE003853; NID
 A:Experimental source: serogroup O1; strain N16961; biotype El Tor
 C:Genetics:
 A:Gene: VCA0218
 A:Map position: 2
 C:Superfamily: phosphatidylcholine-sterol O-acyltransferase

Query Match 48.9%; Score 45; DB 2; Length 418; Best Local Similarity 40.0%; Pred. No. 16; Mismatches 6; Conservative 8; Indels 6; Gaps 0;

Query 1 QOIQTLSKIVARNESLKQ 20
 Db 301 EQIDVRSKIIQGNAPFIREQ 320

RESULT 8
 D82743
 hypothetical protein XF0945 [imported] - Xylella fastidiosa (strain 9a5c)
 C:Species: Xylella fastidiosa
 C:Date: 18-Aug-2000 #sequence_revision 20-Aug-2000 #text_change 09-Jul-2004
 C:Accession: D82743
 R:anonymous, The Xylella fastidiosa Consortium of the Organization for Nucleotide Sequencing
 Nature 406, 151-157, 2000
 A:Title: The genome sequence of the plant pathogen Xylella fastidiosa.
 A:Reference number: A82515; MUID:20365717; PMID:10910347
 A:Note: for a complete list of authors see reference number A59328 below
 A:Accession: D82743
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-71 <SIM>
 A:Cross-references: UNIPROT:Q9PFT3; UNIPARC:UPI00000C2596; GB:AE003933; GB:AE003849; NID
 A:Experimental source: strain 9a5c
 R:Simpton, A.J.G.; Reinach, F.C.; Arruda, P.; Abreu, F.A.; Acencio, M.; Alvarenga, R.; Briones, M.R.S.; Bueno, M.R.P.; Camargo, A.A.; Camargo, L.E.A.; Carraro, D.M.; Carrer, H. as-Neto, E.; Docena, C.; El-Dorri, H.; Facincani, A.P.; Ferreira, A.J.S.
 submitted to GenBank, June 2000
 A:Authors: Ferreira, V.C.A.; Ferro, J.A.; Fraga, J.S.; Franca, S.C.; Frañco, M.C.; Frohm J.D.; Junqueira, M.L.; Kemper, E.L.; Kitajima, J.P.; Krieger, J.E.; Kuranae, E.E.; Laigh chado, M.A.; Madeira, A.M.B.N.; Madeira, H.M.F.; Marino, C.L.; Marques, M.V.; Martins, E. A:Authors: Martins, E.M.F.; Matukuma, A.Y.; Menck, C.F.M.; Miracca, E.C.; Miyaki, C.Y.; F.G.; Nunes, L.R.; Oliveira, M.A.; de Oliveira, M.C.; de Oliveira, R.C.; Palmieri, D.A. Rodrigues, V.; Rosa, A.J. de M.; de Rosa Jr., V.E.; de Sa, R.G.; Santelli, R.V.; Sawasak A:Authors: da Silva, A.C.R.; da Silva, F.R.; da Silva, A.M.; Silva Jr., W.A.; da Silveir M.; Tshako, M.H.; Vallada, H.; Van Sluys, M.A.; Verjovski-Almeida, S.; Vettore, A.L.; Z A:Reference number: A59328
 A:Contents: annotation
 C:Genetics:
 A:Gene: XF0945

Query Match 47.8%; Score 44; DB 2; Length 71; Best Local Similarity 56.2%; Pred. No. 3.4; Mismatches 9; Conservative 3; Mismatches 4; Indels 0; Gaps 0;

Query 1 QOIQTLSKIVARNES 16
 Db 5 QOSQTASRTARNSN 20

RESULT 9
 H72670
 hypothetical protein APE0790 - Aeropyrum pernix (strain K1)
 C:Species: Aeropyrum pernix
 C:Date: 20-Aug-1999 #sequence_revision 20-Aug-1999 #text_change 09-Jul-2004
 C:Accession: H72670
 R:Kawarabayasi, Y.; Hino, Y.; Horikawa, H.; Yamazaki, S.; Haikawa, Y.; Jin-no, K.; Takai awa, H.; Takamiya, M.; Masuda, S.; Funahashi, T.; Tanaka, T.; Kudoh, Y.; Yamazaki, J.; DNA Res. 6, 83-101, 1999
 A:Title: Complete genome sequence of an aerobic hyper-thermophilic Crenarchaeon, Aeropy A:Reference number: A72450; MUID:99310339; PMID:10382966
 A:Accession: H72670
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-297 <KAW>
 A:Cross-references: UNIPROT:Q9YDX9; UNIPARC:UPI000005DCB5; DBJ:AP000060; NID:G5104188; A:Experimental source: strain K1
 C:Genetics:
 A:Gene: APE0790

Query Match 47.8%; Score 44; DB 2; Length 297; Best Local Similarity 47.4%; Pred. No. 16; Mismatches 9; Conservative 5; Mismatches 5; Indels 0; Gaps 0;

Query 1 QOIQTLSKIVARNESLKQ 19
 Db 200 EELATLSKIEYNERIKK 218

RESULT 10
 G83436
 hypothetical protein PA1660 [imported] - Pseudomonas aeruginosa (strain PAO1)
 C:Species: Pseudomonas aeruginosa
 C:Date: 15-Sep-2000 #sequence_revision 15-Sep-2000 #text_change 09-Jul-2004
 C:Accession: G83436
 R:Stover, C.K.; Pham, X.Q.; Erwin, A.L.; Mizoguchi, S.D.; Warren, P.; Hickey, M.J.; B adman, S.; Yuan, Y.; Brody, L.L.; Coulter, S.N.; Folger, K.R.; Kas, A.; Larbig, K.; Lim, .; Lory, S.; Olson, M.V.
 Nature 406, 959-964, 2000
 A:Title: Complete genome sequence of Pseudomonas aeruginosa PAO1, an opportunistic path A:Reference number: A82950; MUID:20437337; PMID:10984043
 A:Accession: G83436
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-526 <STO>
 A:Cross-references: UNIPROT:Q91365; UNIPARC:UPI00000C5432; GB:AE004594; GB:AE004091; NID A:Experimental source: strain PAO1
 C:Genetics:
 A:Gene: PA1660
 C:Superfamily: uncharacterized conserved protein

Query Match 47.8%; Score 44; DB 2; Length 526; Best Local Similarity 40.0%; Pred. No. 30; Mismatches 8; Conservative 7; Mismatches 5; Indels 0; Gaps 0;

Query 1 QOIQTLSKIVARNESLKQ 20
 Db 339 EQIETLSVELTCTQNLPQR 358

RESULT 11
 T15278
 hypothetical protein R155.4 - Caenorhabditis elegans

C;Date: 18-Aug-2000 #sequence_revision 18-Aug-2000 #text_change 09-Jul-2004
C;Accession: T51853
R;Jensen, R.B.; Jensen, K.L.; Jespersen, H.M.; Skriver, K.
FEBS Lett. 436, 283-7, 1998
A;Title: Widespread occurrence of a highly conserved RING-H2 zinc finger motif in the md
A;Reference number: Z13771; MUID:98452956; PMID:9781696
A;Accession: T51853
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: mRNA
A;Residues: 1-329 <JEN>
A;Cross-references: UNIPROT:Q9ZT43; UNIPARC:UPI00000A8E52; EMBL:AF079181; PIDN:AAC69855.

Query Match 45.7%; Score 42; DB 2; Length 329;
Best/Local Similarity 50.0%; Pred. NO. 38;
Matches 8; Conservative 4; Mismatches 4; Indels 0; Gaps 0;

QY 4 QTLSSKIVARNESLQK 19
Db 235 QGLKEKLLARNNSVKE 250

Search completed: January 25, 2006, 19:15:20
Job time : 10.1 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2006 Compugen Ltd.

OM protein - protein search, using sw model

Run on: January 25, 2006, 18:48:24 ; Search time 70.3 Seconds
(without alignments)
200.719 Million cell updates/sec

Title: US-10-797-821-10
Perfect score: 92
Sequence: 1 QOIQTLSKIVARNESLKQQ 20

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 2166443 seqs, 705528306 residues

Total number of hits satisfying chosen parameters: 2166443

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : UniProt 05.80:*
1: uniprot_sprot:*
2: uniprot_trembl:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	92	100.0	211	2 Q9ZAS7_STRMU	Q9zas7 streptococc
2	92	100.0	431	2 Q938V0_STRMU	Q938v0 streptococc
3	92	100.0	431	2 Q938V3_STRMU	Q938v3 streptococc
4	92	100.0	431	2 Q9AG98_STRMU	Q9ag98 streptococc
5	92	100.0	431	2 Q8DWM3_STRMU	Q8dwm3 streptococc
6	92	100.0	432	2 Q938V1_STRMU	Q938v1 streptococc
7	92	100.0	432	2 Q938V2_STRMU	Q938v2 streptococc
8	82	89.1	447	2 Q9AKA4_STRAG	Q9aka4 streptococc
9	82	89.1	447	2 Q8E2H1_STRAS	Q8e2h1 streptococc
10	82	89.1	447	2 Q8E7X9_STRAS	Q8e7x9 streptococc
11	64	69.6	398	2 Q5XEL1_STRP6	Q5xel1 streptococc
12	64	69.6	398	2 Q9ALZ8_STRPY	Q9alz8 streptococc
13	64	69.6	398	2 Q7CNQ7_STRP8	Q7cnq7 streptococc
14	64	69.6	398	2 Q8P318_STRP3	Q8p318 streptococc
15	59	64.1	474	2 Q5M6K4_STRT2	Q5m6k4 streptococc
16	59	64.1	485	2 Q5M212_STRT1	Q5m212 streptococc
17	49.5	53.8	589	2 Q4RY58_TETNG	Q4ry58 tetraodon n
18	49	53.3	392	2 Q8DMY4_STRR6	Q8dm4 streptococc
19	49	53.3	392	2 Q97N55_STRPN	Q97n55 streptococc
20	49	53.3	407	2 Q9VMT7_DROME	Q9vmt7 drosophila
21	49	53.3	7500	2 Q7KTP3_DROME	Q7ktp3 drosophila
22	49	53.3	7735	2 Q9VMT8_DROME	Q9vmt8 drosophila
23	48	52.2	130	2 Q4ZTL1_PSESY	Q4ztl1 pseudomonas
24	48	52.2	206	2 Q6Z912_RAT	Q6z912 rattus norv
25	48	52.2	664	2 Q7M809_WOLUSU	Q7m809 wolinnella s
26	47	51.1	449	2 Q93LK3_ENTFA	Q93lk3 enterococcu
27	47	51.1	470	2 Q87325_VIBMI	Q87325 vibrio mimi
28	46	50.0	397	2 Q5ADL6_CANAL	Q5adl6 candida alb
29	46	50.0	538	2 Q9VBX7_DROME	Q9vbx7 drosophila
30	46	50.0	1186	1 ERCC5_HUMAN	P28715 homo sapien
31	46	50.0	1186	2 Q5JUS5_HUMAN	Q5jus5 homo sapien

RESULT 1
Q9ZAS7_STRMU
ID Q9ZAS7_STRMU PRELIMINARY; PRT; 211 AA.
AC Q9ZAS7;
DT 01-MAY-1999 (TREMBLrel. 10, Created)
DT 01-MAY-1999 (TREMBLrel. 10, Last sequence update)
DT 01-OCT-2002 (TREMBLrel. 22, Last annotation update)
DE Putative secreted protein (fragment).
OS Streptococcus mutans.
OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
OC Streptococcus.
OX NCBI_TaxID=1309;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=GS-5 Kuramitsu;
RA Peruzzi F., Piggot P.J., Daneo-Moore L.;
RL Submitted (NOV-1996) to the EMBL/GenBank/DBJ databases.
DR EMBL; U78607; AAD00288.1; -; Genomic_DNA.
FT NON_TER 211 211
SQ SEQUENCE 211 AA; 22803 MW; 4ACE331159CFATC6 CRC64;

Query Match 100.0%; Score 92; DB 2; Length 211;
Best Local Similarity 100.0%; Pred. No. 1.9e-06;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 QOIQTLSKIVARNESLKQQ 20
|||
DB 88 QOIQTLSKIVARNESLKQQ 107
|||

RESULT 2
Q938V0_STRMU
ID Q938V0_STRMU PRELIMINARY; PRT; 431 AA.
AC Q938V0;
DT 01-DEC-2001 (TREMBLrel. 19, Created)
DT 01-DEC-2001 (TREMBLrel. 19, Last sequence update)
DT 01-MAR-2004 (TREMBLrel. 26, Last annotation update)
DE Glucan-binding protein B.
OS Streptococcus mutans.
OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
OC Streptococcus.
OX NCBI_TaxID=1309;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=5SM3;
RA Jin S., Duncan M.J., Taubman M.A., Smith D.J.;
RT "Cloning of the gopB gene from Streptococcus mutans.";
RL J. Dent. Res. 79:224-224(2000).
RN [2]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=5SM3;
RX MEDLINE=21481971; PubMed=11598068;
RX DOI=10.1128/JAI.69.11.6931-6941.2001;
RA Mattos-Graner R.O., Jin S., King W.F., Chen T., Smith D.J.,

32 46 50.0 1277 2 Q4JF57_TETTH Q4jf57 tetrahymena
33 46 50.0 1611 2 Q59FZ7_HUMAN Q59fz7 homo sapien
34 46 50.0 1677 2 Q7KT48_DROME Q7kt48 drosophila
35 46 50.0 1690 1 C190_DROME Q9vj65 drosophila
36 45.5 49.5 173 2 Q9IQY5_9VIRU Q9iqy5 chinese whe
37 45.5 49.5 173 2 Q9WIE7_9VIRU Q9wie7 chinese whe
38 45 48.9 169 2 Q9AZU7_9CAUD Q9azu7 bacterioph
39 45 48.9 169 2 Q9CFQ3_LACLA Q9cfq3 lactococcus
40 45 48.9 418 2 Q56631_VIBCH Q56631 vibrio chol
41 45 48.9 418 2 Q9KMW0_VIBCH Q9kmw0 vibrio chol
42 45 48.9 538 2 Q8IGB0_DROME Q8igb0 drosophila
43 45 48.9 745 2 Q6BYL6_DEBHA Q6byl6 debaryomyce
44 45 48.9 983 2 Q5GSY1_WOLTR Q5gsy1 wolbachia s
45 45 48.9 2779 1 LVA_DROME Q8ms61 drosophila

ALIGNMENTS

RA Duncan M.J.;
 RT "Cloning of the Streptococcus mutans gene encoding glucan binding
 RT protein B and analysis of genetic diversity and protein production in
 RT clinical isolates.";
 RL Infect. Immun. 69:6931-6941(2001).
 DR EMBL; AY046414; AAK94504.1; -; Genomic_DNA.
 DR InterPro; IPR007921; CHAP.
 DR InterPro; IPR009148; Siba.
 DR Pfam; PF05257; CHAP; 1.
 DR PRINTS; PR01852; SIBAPROTEIN.
 DR PROSITE; PS0911; CHAP; 1.
 SQ SEQUENCE 431 AA; 44650 MW; 05D38D8D8BC4609F CRC64;

Query Match 100.0%; Score 92; DB 2; Length 431;
 Best Local Similarity 100.0%; Pred. No. 4e-06;
 Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 QOIQTLSKIVARNESLKQ 20
 |||||
 Db 88 QOIQTLSKIVARNESLKQ 107

RESULT 3

ID Q938V3_STRMU PRELIMINARY; PRT; 431 AA.
 AC Q938V3;
 DT 01-DEC-2001 (TrEMBLrel. 19, Created)
 DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
 DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
 DE Glucan-binding protein B.
 OS Streptococcus mutans.
 OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
 OC Streptococcus.
 OX NCBI_TaxID=1309;
 RN NUCLEOTIDE SEQUENCE.
 RC STRAIN=3VF4;
 RC STRAIN=5J32;
 RA Jin S., Duncan M.J., Taubman M.A., Smith D.J.;
 RT "Cloning of the gbpB gene from Streptococcus mutans.";
 RL J. Dent. Res. 79:224-224(2000).
 RN [2]
 RP NUCLEOTIDE SEQUENCE.
 RC STRAIN=5J32;
 RC STRAIN=5J32;
 RX MEDLINE=21481971; PubMed=11598068;
 RX DOI=10.1128/IAI.69.11.6931-6941.2001;
 RA Mattos-Graner R.O., Jin S., King W.F., Chen T., Smith D.J.,
 RA Duncan M.J.;
 RT "Cloning of the Streptococcus mutans gene encoding glucan binding
 RT protein B and analysis of genetic diversity and protein production in
 RT clinical isolates.";
 RL Infect. Immun. 69:6931-6941(2001).
 DR EMBL; AY046410; AAK94500.1; -; Genomic_DNA.
 DR InterPro; IPR007921; CHAP.
 DR InterPro; IPR009148; Siba.
 DR Pfam; PF05257; CHAP; 1.
 DR PRINTS; PR01852; SIBAPROTEIN.
 DR PROSITE; PS0911; CHAP; 1.
 SQ SEQUENCE 431 AA; 44620 MW; 464FE3B563FB7E51 CRC64;

Query Match 100.0%; Score 92; DB 2; Length 431;
 Best Local Similarity 100.0%; Pred. No. 4e-06;
 Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 QOIQTLSKIVARNESLKQ 20
 |||||
 Db 88 QOIQTLSKIVARNESLKQ 107

RESULT 4

ID Q9AG98_STRMU PRELIMINARY; PRT; 431 AA.
 AC Q9AG98;
 DT 01-JUN-2001 (TrEMBLrel. 17, Created)

DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
 DT 01-FEB-2005 (TrEMBLrel. 29, Last annotation update)
 DE Immunodominant glycoprotein IDG-60 (Glucan-binding protein B).
 GN Name=sagA;
 OS Streptococcus mutans.
 OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
 OC Streptococcus.
 OX NCBI_TaxID=1309;
 RN [1]
 RP NUCLEOTIDE SEQUENCE.
 RC STRAIN=GS-5;
 RX MEDLINE=21153617; PubMed=11254612;
 RX DOI=10.1128/IAI.69.4.2493-2501.2001;
 RA Chia J.S., Lee Y.Y., Huang P.T., Chen J.Y.;
 RT "Identification of stress-responsive genes in Streptococcus mutans by
 RT differential display reverse transcription-PCR.";
 RL Infect. Immun. 69:2493-2501(2001).
 RN [2]
 RP NUCLEOTIDE SEQUENCE.
 RC STRAIN=GS-5;
 RX MEDLINE=21481977; PubMed=11598074;
 RX DOI=10.1128/IAI.69.11.6987-6998.2001;
 RA Chia J.S., Chang L.Y., Shun C.T., Chang Y.Y., Chen J.Y.;
 RT "A 60-kilodalton immunodominant glycoprotein is essential for cell
 RT wall integrity and the maintenance of cell shape in Streptococcus
 RT mutans.";
 RL Infect. Immun. 69:6987-6998(2001).
 RN [3]
 RP NUCLEOTIDE SEQUENCE.
 RC STRAIN=GS-5;
 RX MEDLINE=21481977; PubMed=11598074;
 RX DOI=10.1128/IAI.69.11.6987-6998.2001;
 RA Chia J.S., Chang L.Y., Shun C.T., Chang Y.Y., Chen J.Y.;
 RT "A 60-kilodalton immunodominant glycoprotein is essential for cell
 RT wall integrity and the maintenance of cell shape in Streptococcus
 RT mutans.";
 RL Infect. Immun. 69:6987-6998(2001).
 RN [4]
 RP NUCLEOTIDE SEQUENCE.
 RC STRAIN=3VF4;
 RC STRAIN=3VF4;
 RA Jin S., Duncan M.J., Taubman M.A., Smith D.J.;
 RT "Cloning of the gbpB gene from Streptococcus mutans.";
 RL J. Dent. Res. 79:224-224(2000).
 RN [5]
 RP NUCLEOTIDE SEQUENCE.
 RC STRAIN=3VF4;
 RC MEDLINE=21481971; PubMed=11598068;
 RX DOI=10.1128/IAI.69.11.6931-6941.2001;
 RA Mattos-Graner R.O., Jin S., King W.F., Chen T., Smith D.J.,
 RA Duncan M.J.;
 RT "Cloning of the Streptococcus mutans gene encoding glucan binding
 RT protein B and analysis of genetic diversity and protein production in
 RT clinical isolates.";
 RL Infect. Immun. 69:6931-6941(2001).
 DR EMBL; AF338445; AAK08104.1; -; Genomic_DNA.
 DR EMBL; AY046411; AAK94501.1; -; Genomic_DNA.
 DR InterPro; IPR007921; CHAP.
 DR InterPro; IPR009148; Siba.
 DR Pfam; PF05257; CHAP; 1.
 DR PRINTS; PR01852; SIBAPROTEIN.
 DR PROSITE; PS0911; CHAP; 1.
 SQ SEQUENCE 431 AA; 44592 MW; 3EBE21FC5E47232E CRC64;

Query Match 100.0%; Score 92; DB 2; Length 431;
 Best Local Similarity 100.0%; Pred. No. 4e-06;
 Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 QOIQTLSKIVARNESLKQ 20
 |||||
 Db 88 QOIQTLSKIVARNESLKQ 107

RESULT 5

ID Q8DWM3_STRMU PRELIMINARY; PRT; 431 AA.
 AC Q8DWM3;
 DT 01-MAR-2003 (TrEMBLrel. 23, Created)
 DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)

DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
 DE Putative secreted antigen GbpB/Saga; putative peptidoglycan
 DE hydrolase.
 GN Name=gbpB; OrderedLocusNames=SMU.22;
 OS Streptococcus mutans.
 OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
 OC Streptococcus.
 OX NCBI_TaxID=1309;
 RN [1]
 RP NUCLEOTIDE SEQUENCE.
 RC STRAIN=UAI59 / ATCC 700610 / Serotype c;
 RX MEDLINE=22295063; PubMed=12397186; DOI=10.1073/pnas.172501299;
 RA Ajdic D.J., McShan W.M., McLaughlin R.E., Savic G., Chang J.,
 RA Carson M.B., Primeaux C., Tian R., Kenton S., Jia H.G., Lin S.P.,
 RA Qian Y., Li S., Zhu H., Najjar F.Z., Lai H., White J., Roe B.A.,
 RA Ferretti J.J.;
 RT "Genome sequence of Streptococcus mutans UAI59, a cariogenic dental
 RT pathogen.";
 RL Proc. Natl. Acad. Sci. U.S.A. 99:14434-14439(2002).
 DR EMBL; AR014855; AAN57811.1; -; Genomic_DNA.
 DR InterPro; IPR007921; CHAP.
 DR InterPro; IPR009148; Siba.
 DR Pfam; PF05257; CHAP; 1.
 DR PRINTS; PR01852; SIBAPROTEIN.
 DR PROSITE; PS50911; CHAP; 1.
 KW Complete proteome.
 SQ SEQUENCE 431 AA; 44620 MW; 2D1CA685248CCD3E CRC64;

 Query Match 100.0%; Score 92; DB 2; Length 431;
 Best Local Similarity 100.0%; Pred. No. 4e-06;
 Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

 QY 1 QOIQTLLSSKIVARNESLKKQ 20
 DB 88 QOIQTLLSSKIVARNESLKKQ 107

 RESULT 6
 Q938V1 STRMU
 ID Q938V1 STRMU PRELIMINARY; PRT; 432 AA.
 AC Q938V1;
 DT 01-DEC-2001 (TrEMBLrel. 19, Created)
 DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
 DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
 DE Glucan-binding protein B.
 OS Streptococcus mutans.
 OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
 OC Streptococcus.
 OX NCBI_TaxID=1309;
 RN [1]
 RP NUCLEOTIDE SEQUENCE.
 RC STRAIN=3SN1;
 RA Jin S., Duncan M.J., Taubman M.A., Smith D.J.;
 RT "Cloning of the gbpB gene from Streptococcus mutans.";
 RL J. Dent. Res. 79:224-224(2000).
 RN [2]
 RP NUCLEOTIDE SEQUENCE.
 RC STRAIN=3SN1;
 RX MEDLINE=21481971; PubMed=11598068;
 RX DOI=10.1128/IAI.69.11.6931-6941.2001;
 RA Mattos-Graner R.O., Jin S., King W.F., Chen T., Smith D.J.,
 RA Duncan M.J.;
 RT "Cloning of the Streptococcus mutans gene encoding glucan binding
 RT protein B and analysis of genetic diversity and protein production in
 RT clinical isolates.";
 RL Infect. Immun. 69:6931-6941(2001).
 DR EMBL; AY046413; AAK94503.1; -; Genomic_DNA.
 DR InterPro; IPR007921; CHAP.
 DR InterPro; IPR009148; Siba.
 DR Pfam; PF05257; CHAP; 1.
 DR PRINTS; PR01852; SIBAPROTEIN.
 DR PROSITE; PS50911; CHAP; 1.
 SQ SEQUENCE 432 AA; 44652 MW; 3F88ECB9A1F3BE4F CRC64;

Query Match 100.0%; Score 92; DB 2; Length 432;
 Best Local Similarity 100.0%; Pred. No. 4.1e-06;
 Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

 QY 1 QOIQTLLSSKIVARNESLKKQ 20
 DB 88 QOIQTLLSSKIVARNESLKKQ 107

 RESULT 7
 Q938V2 STRMU
 ID Q938V2 STRMU PRELIMINARY; PRT; 432 AA.
 AC Q938V2;
 DT 01-DEC-2001 (TrEMBLrel. 19, Created)
 DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
 DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
 DE Glucan-binding protein B.
 OS Streptococcus mutans.
 OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
 OC Streptococcus.
 OX NCBI_TaxID=1309;
 RN [1]
 RP NUCLEOTIDE SEQUENCE.
 RC STRAIN=15JP2;
 RA Jin S., Duncan M.J., Taubman M.A., Smith D.J.;
 RT "Cloning of the gbpB gene from Streptococcus mutans.";
 RL J. Dent. Res. 79:224-224(2000).
 RN [2]
 RP NUCLEOTIDE SEQUENCE.
 RC STRAIN=15JP2;
 RX MEDLINE=21481971; PubMed=11598068;
 RX DOI=10.1128/IAI.69.11.6931-6941.2001;
 RA Mattos-Graner R.O., Jin S., King W.F., Chen T., Smith D.J.,
 RA Duncan M.J.;
 RT "Cloning of the Streptococcus mutans gene encoding glucan binding
 RT protein B and analysis of genetic diversity and protein production in
 RT clinical isolates.";
 RL Infect. Immun. 69:6931-6941(2001).
 DR EMBL; AY046413; AAK94502.1; -; Genomic_DNA.
 DR InterPro; IPR007921; CHAP.
 DR InterPro; IPR009148; Siba.
 DR Pfam; PF05257; CHAP; 1.
 DR PRINTS; PR01852; SIBAPROTEIN.
 DR PROSITE; PS50911; CHAP; 1.
 SQ SEQUENCE 432 AA; 44648 MW; E769B2504AEE50E9 CRC64;

 Query Match 100.0%; Score 92; DB 2; Length 432;
 Best Local Similarity 100.0%; Pred. No. 4.1e-06;
 Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

 QY 1 QOIQTLLSSKIVARNESLKKQ 20
 DB 88 QOIQTLLSSKIVARNESLKKQ 107

 RESULT 8
 Q9AKA4 STRAG
 ID Q9AKA4 STRAG PRELIMINARY; PRT; 447 AA.
 AC Q9AKA4;
 DT 01-JUN-2001 (TrEMBLrel. 17, Created)
 DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
 DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
 DE PcsB protein precursor.
 GN Name=pcsB;
 OS Streptococcus agalactiae.
 OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
 OC Streptococcus.
 OX NCBI_TaxID=1311;
 RN [1]
 RP NUCLEOTIDE SEQUENCE.
 RC STRAIN=6313;
 RX MEDLINE=21101799; PubMed=11157929;

4

[1]
RN RP NUCLEOTIDE SEQUENCE.
RC STRAIN=MGAS8232 / Serotype M18;
RX MEDLINE=21927593; PubMed=11917108; DOI=10.1073/pnas.062526099;
RA SMOOT J.C., Barbican K.D., Van Gompel J.J., Smoot L.M., Chaussee M.S.,
EA Sylva G.L., Sturdevant D.E., Ricklefs S.M., Porcella S.F., Zhang Q.,
RA Parkins L.B., Beres S.B., Campbell D.S., Smith T.M., Zhang J.M.;
RA Kapur V., Daly J.A., Veasy L.G., Musser J.M.;
RT "Genome sequence and comparative microarray analysis of serotype M18
group A Streptococcus strains associated with acute rheumatic fever
outbreaks";
RL Proc. Natl. Acad. Sci. U.S.A. 99:4668-4673(2002).
DR EMBL; AE009955; AAJ96849.1; -; Genomic_DNA.
DR InterPro; IPR007921; CHAP.
DR InterPro; IPR009148; SIBA.
DR Pfam; PF05257; CHAP; 1.
DR PRINTS; PR01852; SIBAPROTEIN.
DR PROSITE; PS50911; CHAP; 1.
RW Complete proteome.
SQ SEQUENCE 398 AA; 42028 MW; 5DCEDA78CB863B60 CRC64;

Query Match 69.6%; Score 64; DB 2; Length 398;
Best Local Similarity 65.0%; Pred No. 0.12; Length 398;
Matches 13; Conservative 4; Mismatches 3; Indels 0; Gaps

Gy 1 QQIQTLLSSKIVARNESLKQQ 20
 | : | : ||||| | | |
Db 85 QBKALTSQIVARNEKLKNQ 104

RESULT 14
QBSP318_STRP3 PRELIMINARY; PRT; 398 AA.

ID QBSP318_STRP3 PRELIMINARY; PRT; 398 AA.
AC QBSP318; Q7CFL7;
DT 01-OCT-2002 (TrEMBLrel. 22, Created)
DT 01-OCT-2002 (TrEMBLrel. 22, Last sequence update)
DT 13-SEP-2005 (TrEMBLrel. 31, Last annotation update)
DE Putative secreted protein.
GN OrderedLocusNames=Sps0015_SpyM3_0014;
OS Streptococcus pyogenes (serotype_M3).
OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
OC Streptococcus.
OX NCBI_TaxID=301448;
RN [1]

RN NP NUCLEOTIDE SEQUENCE.
RC STRAIN=SSI-1 / Serotype M3;
RX MEDLINE=22683278; PubMed=12799345; DOI=10.1101/gr.1096703;
RA Nakagawa I., Kuwakata S., Yamashita A., Nakata M., Tomiyasu Y.,
RA Okahashi N., Kawabata K., Hamada K., Shibata T., Yasunaga T.,
RA Hayashi H., Hattori M., Hanada S.;
RT "Genome sequence of an M3 strain of Streptococcus pyogenes reveals a large-scale genomic rearrangement in invasive strains and new insights into phage evolution.";
RL Genome Res. 13:1042-1055(2003).
RN [2]

RN NP NUCLEOTIDE SEQUENCE.
RC STRAIN=MGAS315 / Serotype M3;
RX MEDLINE=22133808; PubMed=12122206; DOI=10.1073/pnas.152298499;
EA Beres S.B., Sylva G.L., Barbican K.D., Lei B., Hoff J.S.,
RA Mammarella N.D., Liu M.-Y., Smoot J.C., Porcella S.F., Parkins L.D.,
RA Campbell D.S., Smith T.M., McCormick J.K., Leung D.Y.M.,
RA Schlievert P.M., Musser J.M.;
RT "genome sequence of a serotype M3 strain of group A Streptococcus: phage-encoded toxins, the high-virulence phenotype, and clone emergence.";
RL Proc. Natl. Acad. Sci. U.S.A. 99:10078-10083(2002).
DR EMBL; BA000034; BAC63110.1; -; Genomic DNA.
DR ENBL; AE014136; AMW78621.1; -; Genomic_DNA.
DR InterPro; IPR007921; CHAP.
DR InterPro; IPR009148; SIBA.
DR Pfam; PF05257; CHAP; 1.
DR PRINTS; PR01852; SIBAPROTEIN.
DR PROSITE; PS50911; CHAP; 1.

```
KW Complete proteome.
SQ SEQUENCE 398 AA; 42028 MW; 5DCEDA78CB863B60 CRC64;

Query Match          69.6%; Score 64; DB 2; Length 398;
Best Local Similarity 65.0%; Pred. No. 0.12;
Matches 13; Conservative 4; Mismatches 3; Indels 0; Gaps 0;

QY 1 QOIOTLSSKIVARNESLKQK 20
   |:|:|:|:|:|:|:|:|
Db 85 QEIKALTSQIVARNEKLNQ 104

RESULT 15
Q5M6K4_STRT2
ID Q5M6K4_STRT2 PRELIMINARY; PRT; 474 AA.
AC Q5M6K4;
DT 01-FEB-2005 (TrEMBLrel. 29, Created)
DT 01-FEB-2005 (TrEMBLrel. 29, Last sequence update)
DT 13-SEP-2005 (TrEMBLrel. 31, Last annotation update)
DE Glucan binding protein [PcsB].
GN Name=pcsB; OrderedLocustNames=stc0022;
OS Streptococcus thermophilus (strain ATCC BAA-250 / IMG 18311).
OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
OC Streptococcus.
OX NCBI_TaxID=264199;
RN [1]
RP NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
RX PubMed=15543133; DOI=10.1038/nbt1034;
RA Bolotin A., Quinquis B., Renault P., Sorokin A., Ehrlich S.D.,
RA Kulakauskas S., Lapidus A., Goltsman E., Mazur M., Pusch G.D.,
RA Fontstein M., Overbeek R., Kyprides N., Purnelle B., Prozzi D.,
RA Ngui K., Masuy D., Hancy F., Burteau S., Boutry M., Delcour J.,
RA Goffeau A., Hols P.;
RT "Complete sequence and comparative genome analysis of the dairy
RT bacterium Streptococcus thermophilus.";
RL Nat. Biotechnol. 22:1554-1558(2004).
RN [2]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=LMG 18311;
RA Borges F., Layec S., Thibessard A., Fernandez A., Gintz B., Hols P.,
RA Decaris B., Leblond-Bourget N.;
RT "cse, a Chimeric and Variable Gene, Encodes an Extracellular Protein
RT Involved in Cellular Segregation in Streptococcus thermophilus.";
RL J. Bacteriol. 187:2737-2746(2005).
DR EMBL; CP000023; AAV59752.1; -; Genomic DNA.
DR EMBL; AY730643; AAW82375.1; -; Genomic DNA.
DR InterPro; IPR007921; CHAP.
DR InterPro; IPR009148; Siba.
DR Pfam; PF05257; CHAP; 1.
DR PRINTS; PR01852; SIBAPROTEIN.
DR PROSITE; PS50911; CHAP; 1.
KW Complete proteome.
SQ SEQUENCE 474 AA; 48142 MW; 1BAC6A9E0A0A200D CRC64;

Query Match          64.1%; Score 59; DB 2; Length 474;
Best Local Similarity 68.4%; Pred. No. 0.92;
Matches 13; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

QY 2 QIOTLSSKIVARNESLKQK 20
   |:|:|:|:|:|:|:|:|
Db 86 EIQELSSKIVARQDSLAKQ 104

Search completed: January 25, 2006, 19:13:42
Job time : 71.3 secs
```

GenCore version 5.1.6
Copyright (c) 1993 - 2006 Compugen Ltd.

OM protein - protein search, using sw model

Run on: January 25, 2006, 18:50:05 ; Search time 16.6 Seconds
(without alignments)
99.609 Million cell updates/sec

Title: US-10-797-821-10
Perfect score: 92
Sequence: 1 QIQTLSKIVARNESLKQQ 20

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 572060 seqs, 82675679 residues

Total number of hits satisfying chosen parameters: 572060

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued Patents AA:*
1: /cgn2_6/ptodata/1/iaa/5.COMB.pep.*
2: /cgn2_6/ptodata/1/iaa/6.COMB.pep.*
3: /cgn2_6/ptodata/1/iaa/H.COMB.pep.*
4: /cgn2_6/ptodata/1/iaa/PCTUS.COMB.pep.*
5: /cgn2_6/ptodata/1/iaa/RE.COMB.pep.*
6: /cgn2_6/ptodata/1/iaa/backfilea1.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	49	53.3	210	2	US-09-222-938A-67
2	49	53.3	392	2	US-09-583-110-4374
3	49	53.3	399	2	US-09-071-035-484
4	47	51.1	422	2	US-09-071-035-484
5	47	51.1	422	2	US-10-206-576-484
6	47	51.1	449	2	US-09-071-035-482
7	47	51.1	449	2	US-10-206-576-482
8	47	51.1	450	2	US-09-134-000C-5714
9	46	50.0	352	2	US-09-502-540-12334
10	46	50.0	368	2	US-09-248-796A-20840
11	45	48.9	213	2	US-09-270-767-34326
12	45	48.9	213	2	US-09-270-767-49543
13	44	47.8	329	2	US-09-489-039A-10276
14	44	47.8	586	2	US-09-252-991A-22689
15	44	47.8	1780	2	US-09-949-016-6899
16	44	47.8	1786	2	US-09-949-016-7880
17	43	46.7	157	2	US-09-270-767-34016
18	43	46.7	157	2	US-09-270-767-49233
19	42	45.7	256	2	US-09-248-796A-21210
20	42	45.7	327	2	US-09-134-000C-5447
21	42	45.7	436	2	US-09-489-039A-10761
22	42	45.7	431	2	US-09-248-796A-18242
23	41	44.6	205	2	US-09-327-983-16
24	41	44.6	205	2	US-09-549-827A-20
25	41	44.6	497	2	US-09-134-000C-5990
26	41	44.6	546	2	US-09-489-039A-9573
27	41	44.6	552	2	US-09-252-991A-24600

28	41	44.6	1044	2	US-09-107-532A-5229	Sequence 5229, Ap
29	41	44.6	1078	2	US-09-248-796A-20284	Sequence 20284, A
30	41	44.6	1196	1	US-08-144-121-4	Sequence 4, Appli
31	41	44.6	1196	1	US-08-735-893-4	Sequence 4, Appli
32	41	44.6	1196	2	US-10-841-139-4	Sequence 4, Appli
33	41	44.6	1765	2	US-09-562-702A-16	Sequence 16, Appl
34	41	44.6	1765	2	US-09-561-818A-16	Sequence 16, Appl
35	41	44.6	1765	2	US-10-037-182-8	Sequence 8, Appli
36	41	44.6	1786	2	US-09-562-702A-14	Sequence 14, Appl
37	41	44.6	1786	2	US-09-561-818A-14	Sequence 14, Appl
38	41	44.6	1786	2	US-09-561-709B-9	Sequence 9, Appli
39	41	44.6	1786	2	US-09-538-092-869	Sequence 869, App
40	41	44.6	1786	2	US-10-037-182-6	Sequence 6, Appli
41	40.5	44.0	579	2	US-09-949-016-10483	Sequence 10483, A
42	40.5	44.0	583	2	US-09-538-092-1071	Sequence 1071, Ap
43	40.5	44.0	1178	2	US-09-489-039A-9944	Sequence 9944, Ap
44	40	43.5	137	2	US-09-134-000C-6189	Sequence 6189, Ap
45	40	43.5	140	2	US-09-252-991A-27959	Sequence 27959, A

ALIGNMENTS

RESULT 1
US-09-222-938A-67
; Sequence 67, Application US/09222938A
; Patent No. 6437108
; GENERAL INFORMATION:
; APPLICANT: Youngman, Philip
; APPLICANT: Pritz, Chrisian
; APPLICANT: Murphy, Christopher
; APPLICANT: Guzman, Luz-Maria
; TITLE OF INVENTION: ESSENTIAL BACTERIAL GENES AND THEIR USE
; FILE REFERENCE: 07334/060001
; CURRENT APPLICATION NUMBER: US/09/222,938A
; CURRENT FILING DATE: 1998-12-30
; NUMBER OF SEQ ID NOS: 102
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 67
; LENGTH: 210
; TYPE: PRT
; ORGANISM: Streptococcus pneumoniae
US-09-222-938A-67

Query Match 53.3%; Score 49; DB 2; Length 210;
Best Local Similarity 52.6%; Pred. No. 1.5;
Matches 10; Conservative 5; Mismatches 4; Indels 0; Gaps 0;

Qy 2 QIQTLSKIVARNESLKQQ 20
Db 89 EITELSKNIVSRNQSLEKQ 107

RESULT 2
US-09-583-110-4374
; Sequence 4374, Application US/09583110
; Patent No. 6699703
; GENERAL INFORMATION:
; APPLICANT: Lynn Doucette-Stamm et al.
; TITLE OF INVENTION: Nucleic Acid and Amino Acid Sequences Relating to Streptococcus
; FILE REFERENCE: PATH00-07A
; CURRENT APPLICATION NUMBER: US/09/583,110
; CURRENT FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: US 09/107,433
; PRIOR FILING DATE: 1998-06-30
; PRIOR APPLICATION NUMBER: US 60/085,131
; PRIOR FILING DATE: 1998-05-12
; PRIOR APPLICATION NUMBER: US 60/051,553
; PRIOR FILING DATE: 1997-07-02
; NUMBER OF SEQ ID NOS: 5322
; SEQ ID NO 4374
; LENGTH: 392

```
; TYPE: PRT
; ORGANISM: Streptococcus pneumoniae
US-09-583-110-4374

Query Match      53.3%; Score 49; DB 2; Length 392;
Best Local Similarity 52.6%; Pred. No. 3.1;
Matches 10; Conservative 5; Mismatches 4; Indels 0; Gaps 0;

QY      2 QIQTLSSKIVARNESLKQ 20
Db      89 EITELSKNIVSRNQSLKQ 107

RESULT 3
US-09-107-433-3230
; Sequence 3230, Application US/09107433
; Patent No. 6800744
; GENERAL INFORMATION:
; APPLICANT: Lynn A Doucette-Stamm and David Bush
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID
; SEQUENCES RELATING TO STREPTOCOCCUS PNEUMONIAE FOR DIAGNOSTIC
; THERAPEUTICS
; NUMBER OF SEQUENCES: 5206
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: GENOME THERAPEUTICS CORPORATION
; STREET: 100 Beaver Street
; CITY: Waltham
; STATE: Massachusetts
; COUNTRY: USA
; ZIP: 02354
; COMPUTER READABLE FORM:
; MEDIUM TYPE: CD-ROM ISO9660
; COMPUTER: <Unknown>
; OPERATING SYSTEM: <Unknown>
; SOFTWARE: <Unknown>
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/107,433
; FILING DATE: 30-Jun-1998
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 60/ 085131
; FILING DATE: May 12, 1998
; APPLICATION NUMBER: 60/051553
; FILING DATE: July 2, 1997
; ATTORNEY/AGENT INFORMATION:
; NAME: Ariniello, Pamela Deneke
; REGISTRATION NUMBER: 40, 489
; REFERENCE/DOCKET NUMBER: GTC-011
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (781)893-5007
; TELEFAX: (781)893-8277
; INFORMATION FOR SEQ ID NO: 3230:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 399 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; HYPOTHETICAL: YES
; ORIGINAL SOURCE:
; ORGANISM: Streptococcus pneumoniae
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (B) LOCATION 1...399
; SEQUENCE DESCRIPTION: SEQ ID NO: 3230:
US-09-107-433-3230

Query Match      53.3%; Score 49; DB 2; Length 399;
Best Local Similarity 52.6%; Pred. No. 3.1;
Matches 10; Conservative 5; Mismatches 4; Indels 0; Gaps 0;

QY      2 QIQTLSSKIVARNESLKQ 20
Db      96 EITELSKNIVSRNQSLKQ 114

RESULT 4
US-09-071-035-484
; Sequence 484, Application US/09071035
; Patent No. 6448043
; GENERAL INFORMATION:
; APPLICANT: Gil H. Choi
; TITLE OF INVENTION: Enterococcus faecalis Polynucleotides and Polypeptides
; NUMBER OF SEQUENCES: 496
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Human Genome Sciences, Inc.
; STREET: 9410 Key West Avenue
; CITY: Rockville
; STATE: Maryland
; COUNTRY: USA
; ZIP: 20850
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.50 inch, 1.4Mb storage
; COMPUTER: HP Vectra 486/33
; OPERATING SYSTEM: MSDOS version 6.2
; SOFTWARE: ASCII Text
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/071,035
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: A. Anders Brookes
; REGISTRATION NUMBER: 36,373
; REFERENCE/DOCKET NUMBER: PB369P2
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (301) 309-8504
; TELEFAX: (301) 309-8512
; INFORMATION FOR SEQ ID NO: 484:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 422 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-09-071-035-484

Query Match      51.1%; Score 47; DB 2; Length 422;
Best Local Similarity 45.0%; Pred. No. 7.3;
Matches 9; Conservative 5; Mismatches 6; Indels 0; Gaps 0;

QY      1 QQIQTLSSKIVARNESLKQ 20
Db      61 QBITNLNQIKRNEAIKQ 80

RESULT 5
US-10-206-576-484
; Sequence 484, Application US/10206576
; Patent No. 6913907
; GENERAL INFORMATION:
; APPLICANT: Choi et al.
; TITLE OF INVENTION: Enterococcus faecalis Polynucleotides and Polypeptides
; NUMBER OF SEQUENCES: 497
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Human Genome Sciences, Inc.
; STREET: 9410 Key West Avenue
; CITY: Rockville
; STATE: Maryland
; COUNTRY: USA
; ZIP: 20850
; COMPUTER READABLE FORM:
; MEDIUM TYPE: CD-R
; COMPUTER: Dell Latitude
; OPERATING SYSTEM: Windows 98
; SOFTWARE: ASCII Text
```

Query Match 51.1%; Score 47; DB 2; Length 422;
Best Local Similarity 45.0%; Pred. No. 7.3;
Matches 9; Conservative 5; Mismatches 6; Indels

1 QQIQTLSSKIVARNESLKQQ 20
|:|:|:|:|:|:|:
61 QEITNLNQRIEKRNEAIKQ 80

GENERAL INFORMATION:
APPLICANT: Gil H. Choi
TITLE OF INVENTION: Enterococcus faecalis Polynucleotides and Polypeptides

GENERAL INFORMATION:
APPLICANT: Gil H. Choi
TITLE OF INVENTION: Enterococcus faecalis Polynucleotides and Polypeptides

ZIP: 20850
 COMPUTER READABLE FORM:
 MEDIUM TYPE: Diskette, 3.50 inch, 1.4Mb storage
 COMPUTER: HP Vectra 486/33
 OPERATING SYSTEM: MSDOS version 6.2
 SOFTWARE: ASCII Text
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/09/071,035

FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER:

FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: A. Anders Brookes
REGISTRATION NUMBER: 36,373
REFERENCE/DOCKET NUMBER: PB369P2
TELECOMMUNICATION INFORMATION:
TELEPHONE: (301) 309-8504
TELEFAX: (301) 309-8512
INFORMATION FOR SEQ ID NO: 482:
SEQUENCE CHARACTERISTICS:
LENGTH: 449 amino acids

```

; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-09-071-035-482

```

Query Match	51.1%	Score 47	DB 2	Length 449
Best Local Similarity	45.0%	Pred. No. 7.8		
Matches 9	Conservative 5	Mismatches 6	Indels 0	Gaps 0

Qy	1	QOIQT	SSKIVARNES	LKQ	20
	:	:	:	:	:
	:	:	:	:	:
Db	88	QBITNL	QRIEKRNEAI	KQ	107

RESULT 7
US-10-206-576-482
; Sequence 482, Application US/10206576

GENERAL INFORMATION:
APPLICANT: Choi et al.
TITLE OF INVENTION: Enterococcus faecalis Polynucleotides and Polypeptides

```

;
; NUMBER OF SEQUENCES: 497
;
; CORRESPONDENCE ADDRESS:
; ADDRESS: Human Genome Sciences, Inc.
;

```

STREET: 9410 Key West Avenue
CITY: Rockville
STATE: Maryland
COUNTRY: USA

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65
66
67
68
69
70
71
72
73
74
75
76
77
78
79
80
81
82
83
84
85
86
87
88
89
90
91
92
93
94
95
96
97
98
99
100
101
102
103
104
105
106
107
108
109
110
111
112
113
114
115
116
117
118
119
120
121
122
123
124
125
126
127
128
129
130
131
132
133
134
135
136
137
138
139
140
141
142
143
144
145
146
147
148
149
150
151
152
153
154
155
156
157
158
159
160
161
162
163
164
165
166
167
168
169
170
171
172
173
174
175
176
177
178
179
180
181
182
183
184
185
186
187
188
189
190
191
192
193
194
195
196
197
198
199
200
201
202
203
204
205
206
207
208
209
210
211
212
213
214
215
216
217
218
219
220
221
222
223
224
225
226
227
228
229
230
231
232
233
234
235
236
237
238
239
240
241
242
243
244
245
246
247
248
249
250
251
252
253
254
255
256
257
258
259
260
261
262
263
264
265
266
267
268
269
270
271
272
273
274
275
276
277
278
279
280
281
282
283
284
285
286
287
288
289
290
291
292
293
294
295
296
297
298
299
300
301
302
303
304
305
306
307
308
309
310
311
312
313
314
315
316
317
318
319
320
321
322
323
324
325
326
327
328
329
330
331
332
333
334
335
336
337
338
339
340
341
342
343
344
345
346
347
348
349
350
351
352
353
354
355
356
357
358
359
360
361
362
363
364
365
366
367
368
369
370
371
372
373
374
375
376
377
378
379
380
381
382
383
384
385
386
387
388
389
390
391
392
393
394
395
396
397
398
399
400
401
402
403
404
405
406
407
408
409
410
411
412
413
414
415
416
417
418
419
420
421
422
423
424
425
426
427
428
429
430
431
432
433
434
435
436
437
438
439
440
441
442
443
444
445
446
447
448
449
450
451
452
453
454
455
456
457
458
459
460
461
462
463
464
465
466
467
468
469
470
471
472
473
474
475
476
477
478
479
480
481
482
483
484
485
486
487
488
489
490
491
492
493
494
495
496
497
498
499
500
501
502
503
504
505
506
507
508
509
510
511
512
513
514
515
516
517
518
519
520
521
522
523
524
525
526
527
528
529
530
531
532
533
534
535
536
537
538
539
540
541
542
543
544
545
546
547
548
549
550
551
552
553
554
555
556
557
558
559
560
561
562
563
564
565
566
567
568
569
570
571
572
573
574
575
576
577
578
579
580
581
582
583
584
585
586
587
588
589
590
591
592
593
594
595
596
597
598
599
600
601
602
603
604
605
606
607
608
609
610
611
612
613
614
615
616
617
618
619
620
621
622
623
624
625
626
627
628
629
630
631
632
633
634
635
636
637
638
639
640
641
642
643
644
645
646
647
648
649
650
651
652
653
654
655
656
657
658
659
660
661
662
663
664
665
666
667
668
669
670
671
672
673
674
675
676
677
678
679
680
681
682
683
684
685
686
687
688
689
690
691
692
693
694
695
696
697
698
699
700
701
702
703
704
705
706
707
708
709
710
711
712
713
714
715
716
717
718
719
720
721
722
723
724
725
726
727
728
729
730
731
732
733
734
735
736
737
738
739
740
741
742
743
744
745
746
747
748
749
750
751
752
753
754
755
756
757
758
759
760
761
762
763
764
765
766
767
768
769
770
771
772
773
774
775
776
777
778
779
780
781
782
783
784
785
786
787
788
789
790
791
792
793
794
795
796
797
798
799
800
801
802
803
804
805
806
807
808
809
810
811
812
813
814
815
816
817
818
819
820
821
822
823
824
825
826
827
828
829
830
831
832
833
834
835
836
837
838
839
840
84

```

;
; COMPUTER: Dell Latitude
; OPERATING SYSTEM: Windows 98
; SOFTWARE: ASCII Text
;

```

[illegible]

CLASSIFICATION: <UNKNOWN>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 09/071,035
FILING DATE: 1998-05-04

;
 ; APPLICATION NUMBER: US 60/046,655
 ;
 ; FILING DATE: 1997-05-16
 ;
 ; APPLICATION NUMBER: US 60/044,031
 ;

;
 ; FILING DATE: 1997-05-06
 ; APPLICATION NUMBER: US 60/066,009
 ; FILING DATE: 1997-11-14
 ;

NAME: Hyman, Mark J.
REGISTRATION NUMBER: 46,789
REFERENCE/DOCKET NUMBER: DB369D1D1

```

/ INFORMATION FOR SEQ ID NO: 482:
/     SEQUENCE CHARACTERISTICS:
/         LENGTH: 449 amino acids
/

```

```

;
TYPE: amino acid
STRANDEDNESS: single
;
TOPOLOGY: linear
;

```

```

; MOLECULE TYPE: protein
; SEQUENCE DESCRIPTION: SEQ ID NO: 482:
US-10-206-576-482

```

Query Match	51.1%	Score 47;	DB 2;	Length 449;
Best Local Similarity	45.0%;	Pred. No. 7.8;		
Matches 9;	Conservative	5;	Mismatches	6;
			Indels	0;
			Gaps	0;

Qy 1 QOIQTLLSKIVARNESLKQ 20
 | : | | : | | | : |
 | : | | : | | | : |

DB 88 QEITNLNQRIEKRNIAIKNQ 107

RESULTS 8
US-09-134-000C-5714

4

```

; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 6899
; LENGTH: 1780
; TYPE: PRT
; ORGANISM: Human
US-09-949-016-6899

Query Match      47.8%  Score 44;  DB 2;  Length 1780;
Best Local Similarity 31.6%  Pred.No 1.2e+02;
Matches 6;  Conservative 9;  Mismatches 4;  Indels 0;  Gaps 0;

Qy      1  QOIQTLLSSKIVARNESLKQ 19
      |::| : ||: :|: :|:
Db      710  QELETSNKKIITQNRKE 728

Search completed: January 25, 2006, 19:18:15
Job time : 16.6 secs

```

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2006 Compugen Ltd.

OM protein - protein search, using sw model

Run on: January 25, 2006, 19:13:51 ; Search time 60 Seconds
(without alignments)
139.276 Million cell updates/sec

Title: US-10-797-821-10
Perfect score: 92
Sequence: 1 QQIQTLSKIVARNESLKQ 20

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1867569 seqs, 417829326 residues

Total number of hits satisfying chosen parameters: 1867569

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published Applications AA.Main:*
1: /cgn2_6/prodata/1/pubpaa/US07_PUBCOMB.pap:*
2: /cgn2_6/prodata/1/pubpaa/US08_PUBCOMB.pap:*
3: /cgn2_6/prodata/1/pubpaa/US09_PUBCOMB.pap:*
4: /cgn2_6/prodata/1/pubpaa/US10A_PUBCOMB.pap:*
5: /cgn2_6/prodata/1/pubpaa/US10B_PUBCOMB.pap:*
6: /cgn2_6/prodata/1/pubpaa/US11_PUBCOMB.pap:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	92	100.0	20	4	US-10-383-930-10
2	92	100.0	20	5	US-10-797-821-10
3	92	100.0	431	4	US-10-383-930-29
4	92	100.0	431	4	US-10-383-930-30
5	92	100.0	431	4	US-10-383-930-33
6	92	100.0	431	5	US-10-797-821-29
7	92	100.0	431	5	US-10-797-821-30
8	92	100.0	431	5	US-10-797-821-33
9	92	100.0	432	4	US-10-383-930-31
10	92	100.0	432	4	US-10-383-930-32
11	92	100.0	432	5	US-10-797-821-31
12	92	100.0	432	5	US-10-797-821-32
13	64	69.6	398	5	US-10-474-792-600
14	49	53.3	210	4	US-10-154-251-67
15	49	53.3	392	5	US-10-472-928-4652
16	49	53.3	399	5	US-10-617-320-3230
17	49	53.3	415	4	US-10-421-654-20
18	49	53.3	415	5	US-10-796-907-20
19	49	53.3	578	6	US-11-097-143-40206
20	47	51.1	422	3	US-09-071-035-484
21	47	51.1	422	4	US-10-206-576-484
22	47	51.1	422	5	US-10-912-362-484
23	47	51.1	449	3	US-09-071-035-482
24	47	51.1	449	4	US-10-206-576-482
25	47	51.1	449	5	US-10-912-362-482
26	46	50.0	424	3	US-09-525-300-1018
27	46	50.0	1186	4	US-10-341-434-51

28	46	50.0	1690	6	US-11-097-143-10224	Sequence 10224, A
29	46	50.0	1690	6	US-11-097-143-10311	Sequence 10311, A
30	45	48.9	352	4	US-10-425-115-304748	Sequence 304748, A
31	45	48.9	539	4	US-10-425-115-332196	Sequence 332196, A
32	45	48.9	2779	6	US-11-097-143-13905	Sequence 13905, A
33	44	47.8	526	4	US-10-389-647-425	Sequence 425, App
34	44	47.8	1059	4	US-10-259-194A-274	Sequence 274, App
35	44	47.8	1074	4	US-10-437-963-102819	Sequence 102819, A
36	43.5	47.3	487	4	US-10-259-194A-268	Sequence 268, App
37	43	46.7	111	4	US-10-437-963-102826	Sequence 102826, A
38	43	46.7	133	4	US-10-097-111-328	Sequence 328, App
39	43	46.7	166	4	US-10-282-122A-50448	Sequence 50448, A
40	43	46.7	548	4	US-10-425-115-332194	Sequence 332194, A
41	43	46.7	645	4	US-10-425-115-332195	Sequence 332195, A
42	43	46.7	649	4	US-10-425-115-332205	Sequence 332205, A
43	43	46.7	1125	6	US-11-097-143-34887	Sequence 34887, A
44	43	46.7	1174	4	US-10-425-115-204285	Sequence 204285, A
45	43	46.7	1754	4	US-10-425-115-204280	Sequence 204280, A

ALIGNMENTS

RESULT 1

US-10-383-930-10
; Sequence 10, Application US/10383930
; Publication No. US20040127400A1
; GENERAL INFORMATION:
; APPLICANT: Smith, Daniel J
; APPLICANT: Taubman, Martin A
; TITLE OF INVENTION: Immunogenicity of Glucan Binding Protein
; FILE REFERENCE: 25669-018
; CURRENT APPLICATION NUMBER: US/10/383,930
; CURRENT FILING DATE: 2003-03-07
; PRIOR APPLICATION NUMBER: 60/402,483
; PRIOR FILING DATE: 2002-08-08
; PRIOR APPLICATION NUMBER: 60/363,209
; PRIOR FILING DATE: 2002-03-07
; NUMBER OF SEQ ID NOS: 41
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 10
; LENGTH: 20
; TYPE: PRT
; ORGANISM: Streptococcus mutans
US-10-383-930-10

Query Match 100.0%; Score 92; DB 4; Length 20;
Best Local Similarity 100.0%; Pred. No. 4.1e-08;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 QQIQTLSKIVARNESLKQ 20
| | | | | | | | | | | | | | | | | | | | | |
Db 1 QQIQTLSKIVARNESLKQ 20

RESULT 2

US-10-797-821-10
; Sequence 10, Application US/10797821
; Publication No. US20050031633A1
; GENERAL INFORMATION:
; APPLICANT: Smith, Daniel J.
; APPLICANT: Taubman, Martin A.
; TITLE OF INVENTION: Glucan Binding Protein and Glycosyltransferase Immunogens
; FILE REFERENCE: 25669-020
; CURRENT APPLICATION NUMBER: US/10/797,821
; CURRENT FILING DATE: 2004-03-09
; PRIOR APPLICATION NUMBER: 10/383,930
; PRIOR FILING DATE: 2003-03-07
; PRIOR APPLICATION NUMBER: 60/363,209
; PRIOR FILING DATE: 2002-03-07
; PRIOR APPLICATION NUMBER: 60/402,483
; PRIOR FILING DATE: 2002-08-08
; PRIOR APPLICATION NUMBER: 09/290,049

```
; PRIOR FILING DATE: 1999-04-12
; PRIOR APPLICATION NUMBER: 60/081,550
; PRIOR FILING DATE: 1998-04-13
; PRIOR APPLICATION NUMBER: 60/115,142
; PRIOR FILING DATE: 1999-01-08
; NUMBER OF SEQ ID NOS: 45
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 10
; LENGTH: 20
; TYPE: PRT
; ORGANISM: Artificial
; FEATURE:
; OTHER INFORMATION: GbpB peptide
US-10-797-821-10
```

```
Query Match 100.0%; Score 92; DB 5; Length 20;
Best Local Similarity 100.0%; Pred. No. 4.1e-08;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY 1 QOIQTLSKIVARNESLKQ 20
| | | | | | | | | | | | | | | | | |
Db 1 QOIQTLSKIVARNESLKQ 20
```

```
RESULT 3
US-10-383-930-29
; Sequence 29, Application US/10383930
; Publication No. US20040127400A1
; GENERAL INFORMATION:
; APPLICANT: Smith, Daniel J
; APPLICANT: Taubman, Martin A
; TITLE OF INVENTION: Immunogenicity of Glucan Binding Protein
; FILE REFERENCE: 25669-018
; CURRENT APPLICATION NUMBER: US/10/383,930
; CURRENT FILING DATE: 2003-03-07
; PRIOR APPLICATION NUMBER: 60/402,483
; PRIOR FILING DATE: 2002-08-08
; PRIOR APPLICATION NUMBER: 60/363,209
; PRIOR FILING DATE: 2002-03-07
; NUMBER OF SEQ ID NOS: 41
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 29
; LENGTH: 431
; TYPE: PRT
; ORGANISM: Streptococcus mutans
US-10-383-930-29
```

```
Query Match 100.0%; Score 92; DB 4; Length 431;
Best Local Similarity 100.0%; Pred. No. 1.4e-06;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY 1 QOIQTLSKIVARNESLKQ 20
| | | | | | | | | | | | | | | | | |
Db 88 QOIQTLSKIVARNESLKQ 107
```

```
RESULT 4
US-10-383-930-30
; Sequence 30, Application US/10383930
; Publication No. US20040127400A1
; GENERAL INFORMATION:
; APPLICANT: Smith, Daniel J
; APPLICANT: Taubman, Martin A
; TITLE OF INVENTION: Immunogenicity of Glucan Binding Protein
; FILE REFERENCE: 25669-018
; CURRENT APPLICATION NUMBER: US/10/383,930
; CURRENT FILING DATE: 2003-03-07
; PRIOR APPLICATION NUMBER: 60/402,483
; PRIOR FILING DATE: 2002-08-08
; PRIOR APPLICATION NUMBER: 60/363,209
; PRIOR FILING DATE: 2002-03-07
; NUMBER OF SEQ ID NOS: 41
; SOFTWARE: PatentIn version 3.2
```

```
; SEQ ID NO 30
; LENGTH: 431
; TYPE: PRT
; ORGANISM: Streptococcus mutans
US-10-383-930-30
```

```
Query Match 100.0%; Score 92; DB 4; Length 431;
Best Local Similarity 100.0%; Pred. No. 1.4e-06;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY 1 QOIQTLSKIVARNESLKQ 20
| | | | | | | | | | | | | | | | | |
Db 88 QOIQTLSKIVARNESLKQ 107
```

```
RESULT 5
US-10-383-930-33
; Sequence 33, Application US/10383930
; Publication No. US20040127400A1
; GENERAL INFORMATION:
; APPLICANT: Smith, Daniel J
; APPLICANT: Taubman, Martin A
; TITLE OF INVENTION: Immunogenicity of Glucan Binding Protein
; FILE REFERENCE: 25669-018
; CURRENT APPLICATION NUMBER: US/10/383,930
; CURRENT FILING DATE: 2003-03-07
; PRIOR APPLICATION NUMBER: 60/402,483
; PRIOR FILING DATE: 2002-08-08
; PRIOR APPLICATION NUMBER: 60/363,209
; PRIOR FILING DATE: 2002-03-07
; NUMBER OF SEQ ID NOS: 41
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 33
; LENGTH: 431
; TYPE: PRT
; ORGANISM: Streptococcus mutans
US-10-383-930-33
```

```
Query Match 100.0%; Score 92; DB 4; Length 431;
Best Local Similarity 100.0%; Pred. No. 1.4e-06;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY 1 QOIQTLSKIVARNESLKQ 20
| | | | | | | | | | | | | | | | | |
Db 88 QOIQTLSKIVARNESLKQ 107
```

```
RESULT 6
US-10-797-821-29
; Sequence 29, Application US/10797821
; Publication No. US20050031633A1
; GENERAL INFORMATION:
; APPLICANT: Smith, Daniel J.
; APPLICANT: Taubman, Martin A.
; TITLE OF INVENTION: Glucan Binding Protein and Glycosyltransferase Immunogens
; FILE REFERENCE: 25669-020
; CURRENT APPLICATION NUMBER: US/10/797,821
; CURRENT FILING DATE: 2004-03-09
; PRIOR APPLICATION NUMBER: 10/383,930
; PRIOR FILING DATE: 2003-03-07
; PRIOR APPLICATION NUMBER: 60/363,209
; PRIOR FILING DATE: 2002-03-07
; PRIOR APPLICATION NUMBER: 60/402,483
; PRIOR FILING DATE: 2002-08-08
; PRIOR APPLICATION NUMBER: 09/290,049
; PRIOR FILING DATE: 1999-04-12
; PRIOR APPLICATION NUMBER: 60/081,550
; PRIOR FILING DATE: 1998-04-13
; PRIOR APPLICATION NUMBER: 60/115,142
; PRIOR FILING DATE: 1999-01-08
; NUMBER OF SEQ ID NOS: 45
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 29
```

```
; LENGTH: 431
; TYPE: PRT
; ORGANISM: Streptococcus mutans
US-10-797-821-29

Query Match          100.0%; Score 92; DB 5; Length 431;
Best Local Similarity 100.0%; Pred. No. 1.4e-06;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 QOIQTLSKIVARNESLKQ 20
   |||||
Db 88 QOIQTLSKIVARNESLKQ 107

RESULT 7
US-10-797-821-30
; Sequence 30, Application US/10797821
; Publication No. US20050031633A1
; GENERAL INFORMATION:
; APPLICANT: Smith, Daniel J.
; APPLICANT: Taubman, Martin A.
; TITLE OF INVENTION: Glucan Binding Protein and Glycosyltransferase Immunogens
; FILE REFERENCE: 25669-020
; CURRENT APPLICATION NUMBER: US/10/797,821
; CURRENT FILING DATE: 2004-03-09
; PRIOR APPLICATION NUMBER: 10/383,930
; PRIOR FILING DATE: 2003-03-07
; PRIOR APPLICATION NUMBER: 60/363,209
; PRIOR FILING DATE: 2002-03-07
; PRIOR APPLICATION NUMBER: 60/402,483
; PRIOR FILING DATE: 2002-08-08
; PRIOR APPLICATION NUMBER: 09/290,049
; PRIOR FILING DATE: 1999-04-12
; PRIOR APPLICATION NUMBER: 60/081,550
; PRIOR FILING DATE: 1998-04-13
; PRIOR APPLICATION NUMBER: 60/115,142
; PRIOR FILING DATE: 1999-01-08
; NUMBER OF SEQ ID NOS: 45
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 30
; LENGTH: 431
; TYPE: PRT
; ORGANISM: Streptococcus mutans
US-10-797-821-30

Query Match          100.0%; Score 92; DB 5; Length 431;
Best Local Similarity 100.0%; Pred. No. 1.4e-06;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 QOIQTLSKIVARNESLKQ 20
   |||||
Db 88 QOIQTLSKIVARNESLKQ 107

RESULT 8
US-10-797-821-33
; Sequence 33, Application US/10797821
; Publication No. US20050031633A1
; GENERAL INFORMATION:
; APPLICANT: Smith, Daniel J.
; APPLICANT: Taubman, Martin A.
; TITLE OF INVENTION: Glucan Binding Protein and Glycosyltransferase Immunogens
; FILE REFERENCE: 25669-020
; CURRENT APPLICATION NUMBER: US/10/797,821
; CURRENT FILING DATE: 2004-03-09
; PRIOR APPLICATION NUMBER: 10/383,930
; PRIOR FILING DATE: 2003-03-07
; PRIOR APPLICATION NUMBER: 60/363,209
; PRIOR FILING DATE: 2002-03-07
; PRIOR APPLICATION NUMBER: 60/402,483
; PRIOR FILING DATE: 2002-08-08
; PRIOR APPLICATION NUMBER: 09/290,049
; PRIOR FILING DATE: 1999-04-12
```

```
; PRIOR APPLICATION NUMBER: 60/081,550
; PRIOR FILING DATE: 1998-04-13
; PRIOR APPLICATION NUMBER: 60/115,142
; PRIOR FILING DATE: 1999-01-08
; NUMBER OF SEQ ID NOS: 45
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 33
; LENGTH: 431
; TYPE: PRT
; ORGANISM: Streptococcus mutans
US-10-797-821-33

Query Match          100.0%; Score 92; DB 5; Length 431;
Best Local Similarity 100.0%; Pred. No. 1.4e-06;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 QOIQTLSKIVARNESLKQ 20
   |||||
Db 88 QOIQTLSKIVARNESLKQ 107

RESULT 9
US-10-383-930-31
; Sequence 31, Application US/10383930
; Publication No. US20040127400A1
; GENERAL INFORMATION:
; APPLICANT: Smith, Daniel J.
; APPLICANT: Taubman, Martin A.
; TITLE OF INVENTION: Immunogenicity of Glucan Binding Protein
; FILE REFERENCE: 25669-018
; CURRENT APPLICATION NUMBER: US/10/383,930
; CURRENT FILING DATE: 2003-03-07
; PRIOR APPLICATION NUMBER: 60/402,483
; PRIOR FILING DATE: 2002-08-08
; PRIOR APPLICATION NUMBER: 60/363,209
; PRIOR FILING DATE: 2002-03-07
; NUMBER OF SEQ ID NOS: 41
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 31
; LENGTH: 432
; TYPE: PRT
; ORGANISM: Streptococcus mutans
US-10-383-930-31

Query Match          100.0%; Score 92; DB 4; Length 432;
Best Local Similarity 100.0%; Pred. No. 1.4e-06;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 QOIQTLSKIVARNESLKQ 20
   |||||
Db 88 QOIQTLSKIVARNESLKQ 107

RESULT 10
US-10-383-930-32
; Sequence 32, Application US/10383930
; Publication No. US20040127400A1
; GENERAL INFORMATION:
; APPLICANT: Smith, Daniel J.
; APPLICANT: Taubman, Martin A.
; TITLE OF INVENTION: Immunogenicity of Glucan Binding Protein
; FILE REFERENCE: 25669-018
; CURRENT APPLICATION NUMBER: US/10/383,930
; CURRENT FILING DATE: 2003-03-07
; PRIOR APPLICATION NUMBER: 60/402,483
; PRIOR FILING DATE: 2002-08-08
; PRIOR APPLICATION NUMBER: 60/363,209
; PRIOR FILING DATE: 2002-03-07
; NUMBER OF SEQ ID NOS: 41
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 32
; LENGTH: 432
; TYPE: PRT
```

```
; ORGANISM: Streptococcus mutans
US-10-383-930-32

Query Match          100.0%; Score 92; DB 4; Length 432;
Best Local Similarity 100.0%; Pred. No. 1.4e-06;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 QOIQTLSKIVARNESLKQQ 20
   |||||
Db 88 QOIQTLSKIVARNESLKQQ 107

RESULT 11
US-10-797-821-31
; Sequence 31, Application US/10797821
; Publication No. US20050031633A1
; GENERAL INFORMATION:
; APPLICANT: Smith, Daniel J.
; TITLE OF INVENTION: Glucan Binding Protein and Glycosyltransferase Immunogens
; FILE REFERENCE: 25669-020
; CURRENT FILING DATE: 2004-03-09
; PRIOR APPLICATION NUMBER: 10/383,930
; PRIOR FILING DATE: 2003-03-07
; PRIOR APPLICATION NUMBER: 60/363,209
; PRIOR FILING DATE: 2002-03-07
; PRIOR APPLICATION NUMBER: 60/402,483
; PRIOR FILING DATE: 2002-08-08
; PRIOR APPLICATION NUMBER: 09/290,049
; PRIOR FILING DATE: 1999-04-12
; PRIOR APPLICATION NUMBER: 60/081,550
; PRIOR FILING DATE: 1998-04-13
; PRIOR APPLICATION NUMBER: 60/115,142
; PRIOR FILING DATE: 1999-01-08
; NUMBER OF SEQ ID NOS: 45
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 31
; LENGTH: 432
; TYPE: PRT
; ORGANISM: Streptococcus mutans
US-10-797-821-31

Query Match          100.0%; Score 92; DB 5; Length 432;
Best Local Similarity 100.0%; Pred. No. 1.4e-06;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 QOIQTLSKIVARNESLKQQ 20
   |||||
Db 88 QOIQTLSKIVARNESLKQQ 107

RESULT 12
US-10-797-821-32
; Sequence 32, Application US/10797821
; Publication No. US20050031633A1
; GENERAL INFORMATION:
; APPLICANT: Smith, Daniel J.
; TITLE OF INVENTION: Glucan Binding Protein and Glycosyltransferase Immunogens
; FILE REFERENCE: 25669-020
; CURRENT FILING DATE: 2004-03-09
; PRIOR APPLICATION NUMBER: 10/383,930
; PRIOR FILING DATE: 2003-03-07
; PRIOR APPLICATION NUMBER: 60/363,209
; PRIOR FILING DATE: 2002-03-07
; PRIOR APPLICATION NUMBER: 60/402,483
; PRIOR FILING DATE: 2002-08-08
; PRIOR APPLICATION NUMBER: 09/290,049
; PRIOR FILING DATE: 1999-04-12
; PRIOR APPLICATION NUMBER: 60/081,550
; PRIOR FILING DATE: 1998-04-13

; ORGANISM: Streptococcus mutans
US-10-383-930-32

Query Match          100.0%; Score 92; DB 4; Length 432;
Best Local Similarity 100.0%; Pred. No. 1.4e-06;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 QOIQTLSKIVARNESLKQQ 20
   |||||
Db 88 QOIQTLSKIVARNESLKQQ 107

RESULT 13
US-10-474-792-600
; Sequence 600, Application US/10474792
; Publication No. US20040236072A1
; GENERAL INFORMATION:
; APPLICANT: Olmsted, Stephen
; APPLICANT: Zagursky, Robert
; APPLICANT: Nickbarg, Elliot
; APPLICANT: Winter, Lourie
; TITLE OF INVENTION: SURFACE PROTEINS OF STREPTOCOCCUS PYOGENES
; FILE REFERENCE: AM 100399
; CURRENT APPLICATION NUMBER: US/10/474,792
; CURRENT FILING DATE: 2003-10-14
; NUMBER OF SEQ ID NOS: 674
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 600
; LENGTH: 398
; TYPE: PRT
; ORGANISM: Streptococcus pyogenes
US-10-474-792-600

Query Match          69.6%; Score 64; DB 5; Length 398;
Best Local Similarity 65.0%; Pred. No. 0.055;
Matches 13; Conservative 4; Mismatches 3; Indels 0; Gaps 0;

QY 1 QOIQTLSKIVARNESLKQQ 20
   |||||
Db 85 QEIKALTSQIVARNEKLKQ 104

RESULT 14
US-10-154-251-67
; Sequence 67, Application US/10154251
; Publication No. US20030092024A1
; GENERAL INFORMATION:
; APPLICANT: Youngman, Philip
; APPLICANT: Fritz, Christian
; APPLICANT: Murphy, Christopher
; APPLICANT: Guzman, Luz-Maria
; TITLE OF INVENTION: ESSENTIAL BACTERIAL GENES AND THEIR USE
; FILE REFERENCE: 06286-060002
; CURRENT APPLICATION NUMBER: US/10/154,251
; CURRENT FILING DATE: 2002-09-16
; PRIOR APPLICATION NUMBER: US 10/154,251
; PRIOR FILING DATE: 2002-05-22
; NUMBER OF SEQ ID NOS: 102
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 67
; LENGTH: 210
; TYPE: PRT
; ORGANISM: Streptococcus pneumoniae
US-10-154-251-67

Query Match          53.3%; Score 49; DB 4; Length 210;
```

```
Qy      2 QIOTLSSKIVARNESLKQQ 20
      :| || ||:|:|:|:|
Db     89 EITELSKNIVSRNQSLKQ 107
```

```

RESULT 15
US-10-472-928-4652
; Sequence 4652, Application US/10472928
; Publication No. US2005020813A1
; GENERAL INFORMATION:
; APPLICANT: CHIRON SCA
; TITLE OF INVENTION: THE INSTITUTE FOR GENOMIC RESEARCH
; FILE REFERENCE: P026926WO
; CURRENT APPLICATION NUMBER: US/10/472,928
; CURRENT FILING DATE: 2003-09-26
; PRIOR APPLICATION NUMBER: GB-0107658.7
; PRIOR FILING DATE: 2001-03-27
; NUMBER OF SEQ ID NOS: 4979
; SOFTWARE: Seqwin99, version 1.03
; SEQ ID NO 4652
; LENGTH: 392
; TYPE: PRT
; ORGANISM: Streptococcus pneumoniae
; FEATURE:
; OTHER INFORMATION: secreted 45 kd protein (usp45)
; OTHER INFORMATION: Cellular location: outside
; OTHER INFORMATION: Similar to strain R6 sequence 15904062 (0.E+01)
US-10-472-928-4652

```

QY 2 QIQTLLSSKI VARNESLKQ 20
:| || ||:|:|:|:|:
Db 89 EITELSKNIVSRNQSLKQ 107

Search completed: January 25, 2006, 20:10:04
Job time : 60 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2006 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: January 25, 2006, 19:15:26 ; Search time 5.5 Seconds
(without alignments)
39.378 Million cell updates/sec

Title: US-10-797-821-10
Perfect score: 92
Sequence: 1 QOQTLSKIVARNESLKQ 20

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 75621 seqs, 10829074 residues

Total number of hits satisfying chosen parameters: 75621

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

- Database : Published Applications AA_New.*
- 1: /cgn2_6/prodata/2/pubpaa/US08_NEW_PUB.pep.*
 - 2: /cgn2_6/prodata/2/pubpaa/US06_NEW_PUB.pep.*
 - 3: /cgn2_6/prodata/2/pubpaa/US07_NEW_PUB.pep.*
 - 4: /cgn2_6/prodata/2/pubpaa/PCT_NEW_PUB.pep.*
 - 5: /cgn2_6/prodata/2/pubpaa/US09_NEW_PUB.pep.*
 - 6: /cgn2_6/prodata/2/pubpaa/US10_NEW_PUB.pep.*
 - 7: /cgn2_6/prodata/2/pubpaa/US11_NEW_PUB.pep.*
 - 8: /cgn2_6/prodata/2/pubpaa/US60_NEW_PUB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
1	92	100.0	431	7	US-11-052-554A-210
2	64	69.6	398	7	US-11-052-554A-252
3	41	44.6	147	6	US-10-821-234-1230
4	40	43.5	252	6	US-10-793-626-1032
5	40	43.5	575	6	US-10-959-310-7
6	40	43.5	575	6	US-10-959-310-8
7	40	43.5	575	7	US-11-131-212-23
8	40	43.5	575	7	US-11-131-212-24
9	40	43.5	766	6	US-10-793-626-420
10	40	43.5	2323	6	US-10-793-626-760
11	39	42.4	171	6	US-10-506-796A-9
12	39	42.4	181	6	US-10-980-388-100
13	39	42.4	342	6	US-10-980-388-118
14	39	42.4	342	7	US-11-174-816-18
15	39	42.4	342	7	US-11-174-819-71
16	38	41.3	124	7	US-11-007-282-18
17	38	41.3	241	7	US-11-152-366-279
18	38	41.3	359	7	US-11-152-366-38
19	38	41.3	738	6	US-10-995-561-692
20	38	41.3	738	6	US-10-995-561-693
21	37.5	40.8	465	6	US-10-878-556A-116
22	37.5	40.8	466	7	US-11-173-672-1
23	37	40.2	136	6	US-10-821-234-1616
24	37	40.2	708	6	US-10-623-155-369
25	37	40.2	1652	6	US-10-995-561-663

26	37	40.2	1938	6	US-10-995-561-661	Sequence 661, App
27	37	40.2	1938	6	US-10-995-561-662	Sequence 662, App
28	37	40.2	1954	6	US-10-995-561-660	Sequence 660, App
29	37	40.2	1972	6	US-10-995-561-664	Sequence 664, App
30	37	40.2	1972	6	US-10-995-561-666	Sequence 666, App
31	37	40.2	2101	6	US-10-857-780-23	Sequence 23, Appl
32	37	40.2	2902	7	US-11-052-554A-91	Sequence 91, Appl
33	36	39.1	99	6	US-10-485-517-299	Sequence 299, App
34	36	39.1	103	7	US-11-007-282-12	Sequence 12, Appl
35	36	39.1	124	7	US-11-007-282-10	Sequence 10, Appl
36	36	39.1	446	7	US-11-119-351-8	Sequence 8, Appl
37	36	39.1	690	6	US-10-613-744-19	Sequence 19, Appl
38	36	39.1	690	6	US-10-511-657-2	Sequence 2, Appl
39	36	39.1	951	7	US-11-121-438-14	Sequence 14, Appl
40	36	39.1	1286	6	US-10-877-346-38	Sequence 38, Appl
41	36	39.1	1410	6	US-10-878-556A-136	Sequence 136, App
42	36	39.1	1586	6	US-10-821-234-901	Sequence 901, App
43	36	39.1	1597	6	US-10-877-346-41	Sequence 41, Appl
44	36	39.1	1613	6	US-10-055-877-145	Sequence 145, App
45	36	39.1	1637	6	US-10-055-877-144	Sequence 144, App

ALIGNMENTS

RESULT 1
US-11-052-554A-210
; Sequence 210, Application US/11052554A
; Publication No. US20050288866A1
; GENERAL INFORMATION:
; APPLICANT: Sachdeva, et al.
; TITLE OF INVENTION: COMPUTATIONAL METHOD FOR IDENTIFYING ADHESIN AND ADHESIN-LIKE
; FILE REFERENCE: 30853/40359A
; CURRENT APPLICATION NUMBER: US/11/052,554A
; CURRENT FILING DATE: 2005-02-07
; PRIOR APPLICATION NUMBER: US 60/589,227
; PRIOR FILING DATE: 2004-07-20
; PRIOR APPLICATION NUMBER: IN 173/DEL/2004
; PRIOR FILING DATE: 2004-02-06
; NUMBER OF SEQ ID NOS: 763
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 210
; LENGTH: 431
; TYPE: PRT
; ORGANISM: Streptococcus mutans UA159
US-11-052-554A-210

Query Match 100.0%; Score 92; DB 7; Length 431;
Best Local Similarity 100.0%; Pred. No. 5.5e-08;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 QOQTLSKIVARNESLKQ 20
Db 88 QOQTLSKIVARNESLKQ 107

RESULT 2
US-11-052-554A-252
; Sequence 252, Application US/11052554A
; Publication No. US20050288866A1
; GENERAL INFORMATION:
; APPLICANT: Sachdeva, et al.
; TITLE OF INVENTION: COMPUTATIONAL METHOD FOR IDENTIFYING ADHESIN AND ADHESIN-LIKE
; FILE REFERENCE: 30853/40359A
; CURRENT APPLICATION NUMBER: US/11/052,554A
; CURRENT FILING DATE: 2005-02-07
; PRIOR APPLICATION NUMBER: US 60/589,227
; PRIOR FILING DATE: 2004-07-20
; PRIOR APPLICATION NUMBER: IN 173/DEL/2004
; PRIOR FILING DATE: 2004-02-06
; NUMBER OF SEQ ID NOS: 763

```

; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 252
; LENGTH: 398
; TYPE: PRT
; ORGANISM: Streptococcus pyogenes MGAS8232
US-11-052-554A-252

Query Match          59.6%; Score 64; DB 7; Length 398;
Best Local Similarity 65.0%; Pred. No. 0.0021; Indels 3; Gaps 0;
Matches 13; Conservative 4; Mismatches 0; Gaps 0;

QY 1 QOIQTLSKIVARNESLKQ 20
   |:|:|:|:|:|:|:|:|:|
Db 85 QEIKALTSQIVARNELKQ 104

RESULT 3
US-10-821-234-1230
; Sequence 1230, Application US/10821234
; Publication No. US20050255114A1
; GENERAL INFORMATION:
; APPLICANT: Labat, Ivan
; APPLICANT: Stache-Crain, Birgit
; APPLICANT: Andarmani, Susan
; APPLICANT: Tang, Y. Tom
; TITLE OF INVENTION: Methods for Diagnosis and Treatment of Preeclampsia
; FILE REFERENCE: 821A
; CURRENT APPLICATION NUMBER: US/10/821,234
; CURRENT FILING DATE: 2004-04-07
; PRIOR APPLICATION NUMBER: US 60/462,047
; PRIOR FILING DATE: 2003-04-07
; NUMBER OF SEQ ID NOS: 1704
; SOFTWARE: pt_seq_genes Version 1.0
; SEQ ID NO 1230
; LENGTH: 147
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-821-234-1230

Query Match          44.6%; Score 41; DB 6; Length 147;
Best Local Similarity 52.6%; Pred. No. 4;
Matches 10; Conservative 2; Mismatches 7; Indels 7; Gaps 0;

QY 1 QOIQTLSKIVARNESLKQ 19
   |:|:|:|:|:|:|:|
Db 51 QOLQLNLTDIRERVESLSQ 69

RESULT 4
US-10-793-626-1032
; Sequence 1032, Application US/10793626
; Publication No. US20050255478A1
; GENERAL INFORMATION:
; APPLICANT: KIMBERLY, WILLIAM JOHN
; TITLE OF INVENTION: STAPHYLOCOCCUS EPIDERMIDIS NUCLEIC ACIDS AND PROTEINS
; FILE REFERENCE: P03480US
; CURRENT APPLICATION NUMBER: US/10/793,626
; CURRENT FILING DATE: 2004-03-04
; PRIOR APPLICATION NUMBER: 60/164,258
; PRIOR FILING DATE: 1999-11-09
; NUMBER OF SEQ ID NOS: 4472
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1032
; LENGTH: 252
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: synthetic
; OTHER INFORMATION: amino acid sequence
US-10-793-626-1032

Query Match          43.5%; Score 40; DB 6; Length 252;
Best Local Similarity 42.1%; Pred. No. 11;
Matches 9; Conservative 2; Mismatches 2; Indels 2; Gaps 0;

QY 8 SKIVARNESLKQ 20
   |:|:|:|:|:|:|:|
Db 44 SKILAKLERLKQ 56

RESULT 5
US-10-959-310-7
; Sequence 7, Application US/10959310
; Publication No. US20050287138A1
; GENERAL INFORMATION:
; APPLICANT: KYOWA HAKKO KOGYO CO., LTD.
; TITLE OF INVENTION: CCR4-specific antibody composition
; FILE REFERENCE: 249-363
; CURRENT APPLICATION NUMBER: US/10/959,310
; CURRENT FILING DATE: 2004-10-07
; PRIOR APPLICATION NUMBER: JP 2003-350162
; PRIOR FILING DATE: 2003-10-08
; PRIOR APPLICATION NUMBER: US 60/572,784
; PRIOR FILING DATE: 2004-05-21
; NUMBER OF SEQ ID NOS: 46
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 7
; LENGTH: 575
; TYPE: PRT
; ORGANISM: Cricetulus griseus
US-10-959-310-7

Query Match          43.5%; Score 40; DB 6; Length 575;
Best Local Similarity 69.2%; Pred. No. 30;
Matches 9; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 8 SKIVARNESLKQ 20
   |:|:|:|:|:|:|:|
Db 44 SKILAKLERLKQ 56

RESULT 6
US-10-959-310-8
; Sequence 8, Application US/10959310
; Publication No. US20050287138A1
; GENERAL INFORMATION:
; APPLICANT: KYOWA HAKKO KOGYO CO., LTD.
; TITLE OF INVENTION: CCR4-specific antibody composition
; FILE REFERENCE: 249-363
; CURRENT APPLICATION NUMBER: US/10/959,310
; CURRENT FILING DATE: 2004-10-07
; PRIOR APPLICATION NUMBER: JP 2003-350162
; PRIOR FILING DATE: 2003-10-08
; PRIOR APPLICATION NUMBER: US 60/572,784
; PRIOR FILING DATE: 2004-05-21
; NUMBER OF SEQ ID NOS: 46
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 8
; LENGTH: 575
; TYPE: PRT
; ORGANISM: Mus musculus
US-10-959-310-8

Query Match          43.5%; Score 40; DB 6; Length 575;
Best Local Similarity 69.2%; Pred. No. 30;
Matches 9; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 8 SKIVARNESLKQ 20
   |:|:|:|:|:|:|:|
Db 44 SKILAKLERLKQ 56

RESULT 7
US-11-131-212-23
; Sequence 23, Application US/11131212
; Publication No. US20050262593A1

```


GENERAL INFORMATION:
; APPLICANT: Yutaka KANDA
; APPLICANT: Mitsuo SATOH
; APPLICANT: Kazuyasu NAKAMURA
; APPLICANT: Kazuhisa UCHIDA
; APPLICANT: Toyohide SHINKAWA
; APPLICANT: Naoko YAMANE
; APPLICANT: Motoo YAMASAKI
; APPLICANT: Nobuo HANAI
; TITLE OF INVENTION: ANTIBODY COMPOSITION-PRODUCING CELL
; FILE REFERENCE: 249-202
; CURRENT APPLICATION NUMBER: US/11/131,212
; CURRENT FILING DATE: 2005-05-18
; PRIOR APPLICATION NUMBER: US/09/971,773
; PRIOR FILING DATE: 2002-08-30
; PRIOR APPLICATION NUMBER: JP 2000-308526
; PRIOR FILING DATE: 2000-10-06
; PRIOR APPLICATION NUMBER: US 60/268,926
; PRIOR FILING DATE: 2001-02-16
; NUMBER OF SEQ ID NOS: 73
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 23
; LENGTH: 575
; TYPE: PRT
; ORGANISM: Cricetulus griseus
US-11-131-212-23

Query Match 43.5%; Score 40; DB 7; Length 575;
Best Local Similarity 69.2%; Pred. No. 30;
Matches 9; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 8 SKIVARNESLKQ 20
DB 44 SKILAKLERLKQ 56

RESULT 8
US-11-131-212-24
; Sequence 24, Application US/11/131212
; Publication No. US20050262593A1
; GENERAL INFORMATION:
; APPLICANT: Yutaka KANDA
; APPLICANT: Mitsuo SATOH
; APPLICANT: Kazuyasu NAKAMURA
; APPLICANT: Kazuhisa UCHIDA
; APPLICANT: Toyohide SHINKAWA
; APPLICANT: Naoko YAMANE
; APPLICANT: Motoo YAMASAKI
; APPLICANT: Nobuo HANAI
; TITLE OF INVENTION: ANTIBODY COMPOSITION-PRODUCING CELL
; FILE REFERENCE: 249-202
; CURRENT APPLICATION NUMBER: US/11/131,212
; CURRENT FILING DATE: 2005-05-18
; PRIOR APPLICATION NUMBER: US/09/971,773
; PRIOR FILING DATE: 2002-08-30
; PRIOR APPLICATION NUMBER: JP 2000-308526
; PRIOR FILING DATE: 2000-10-06
; PRIOR APPLICATION NUMBER: US 60/268,926
; PRIOR FILING DATE: 2001-02-16
; NUMBER OF SEQ ID NOS: 73
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 24
; LENGTH: 575
; TYPE: PRT
; ORGANISM: Mus musculus
US-11-131-212-24

Query Match 43.5%; Score 40; DB 7; Length 575;
Best Local Similarity 69.2%; Pred. No. 30;
Matches 9; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 8 SKIVARNESLKQ 20
DB 44 SKILAKLERLKQ 56

Db 44 SKILAKLERLKQ 56

RESULT 9
US-10-793-626-420
; Sequence 420, Application US/10793626
; Publication No. US20050255478A1
; GENERAL INFORMATION:
; APPLICANT: KIMMERLY, WILLIAM JOHN
; TITLE OF INVENTION: STAPHYLOCOCCUS EPIDERMIDIS NUCLEIC ACIDS AND PROTEINS
; FILE REFERENCE: PU3480US
; CURRENT APPLICATION NUMBER: US/10/793,626
; CURRENT FILING DATE: 2004-03-04
; PRIOR APPLICATION NUMBER: 60/164,258
; PRIOR FILING DATE: 1999-11-09
; NUMBER OF SEQ ID NOS: 4472
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 420
; LENGTH: 766
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: synthetic
; OTHER INFORMATION: amino acid sequence
US-10-793-626-420

Query Match 43.5%; Score 40; DB 6; Length 766;
Best Local Similarity 42.1%; Pred. No. 42;
Matches 8; Conservative 5; Mismatches 6; Indels 0; Gaps 0;

QY 2 QIQTLSKIVARNESLKQ 20
DB 167 QLQRLSKLIERHEILRTQ 185

RESULT 10
US-10-793-626-760
; Sequence 760, Application US/10793626
; Publication No. US20050255478A1
; GENERAL INFORMATION:
; APPLICANT: KIMMERLY, WILLIAM JOHN
; TITLE OF INVENTION: STAPHYLOCOCCUS EPIDERMIDIS NUCLEIC ACIDS AND PROTEINS
; FILE REFERENCE: PU3480US
; CURRENT APPLICATION NUMBER: US/10/793,626
; CURRENT FILING DATE: 2004-03-04
; PRIOR APPLICATION NUMBER: 60/164,258
; PRIOR FILING DATE: 1999-11-09
; NUMBER OF SEQ ID NOS: 4472
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 760
; LENGTH: 2323
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: synthetic
; OTHER INFORMATION: amino acid sequence
US-10-793-626-760

Query Match 43.5%; Score 40; DB 6; Length 2323;
Best Local Similarity 42.1%; Pred. No. 1.6e+02;
Matches 8; Conservative 5; Mismatches 6; Indels 0; Gaps 0;

QY 2 QIQTLSKIVARNESLKQ 20
DB 1020 QLQRLSKLIERHEILRTQ 1038

RESULT 11
US-10-506-796A-9
; Sequence 9, Application US/10506796A
; Publication No. US20060013831A1
; GENERAL INFORMATION:
; APPLICANT: MOR, Tsafir


```
; APPLICANT: LINDEMANN, LOTHAR
; TITLE OF INVENTION: HUMAN TRACE AMINE ASSOCIATED RECEPTORS
; FILE REFERENCE: 22645
; CURRENT APPLICATION NUMBER: US/11/174,816
; CURRENT FILING DATE: 2005-07-05
; PRIOR APPLICATION NUMBER: EP 04103261.6
; PRIOR FILING DATE: 2004-07-08
; NUMBER OF SEQ ID NOS: 115
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 18
; LENGTH: 342
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: hutAAR 8
US-11-174-816-18
```

```
Query Match 42.4%; Score 39; DB 7; Length 342;
Best Local Similarity 47.1%; Pred. No. 23;
Matches 8; Conservative 5; Mismatches 4; Indels 0; Gaps 0;
```

```
QY 2 QIOTLSSKIVARNESLK 18
Db 229 KIETSSKVESSESYK 245
```

```
RESULT 15
US-11-174-819-71
; Sequence 71, Application US/11/174819
; Publication No. US2006000880A1
; GENERAL INFORMATION:
; APPLICANT: EBELING, MARTIN
; APPLICANT: HOENER, MARIUS
; APPLICANT: LINDEMANN, LOTHAR
; TITLE OF INVENTION: MOUSE TRACE AMINE ASSOCIATED RECEPTORS AND
; FILE REFERENCE: 22646
; CURRENT APPLICATION NUMBER: US/11/174,819
; CURRENT FILING DATE: 2005-07-05
; PRIOR APPLICATION NUMBER: EP 04103262.4
; PRIOR FILING DATE: 2004-07-08
; NUMBER OF SEQ ID NOS: 136
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 71
; LENGTH: 342
; TYPE: PRT
; ORGANISM: Homo sapiens
US-11-174-819-71
```

```
Query Match 42.4%; Score 39; DB 7; Length 342;
Best Local Similarity 47.1%; Pred. No. 23;
Matches 8; Conservative 5; Mismatches 4; Indels 0; Gaps 0;
```

```
QY 2 QIOTLSSKIVARNESLK 18
Db 229 KIETSSKVESSESYK 245
```

```
Search completed: January 25, 2006, 20:11:09
Job time : 5.5 secs
```

THIS PAGE BLANK (USPTO)